DISSERTATION ABSTRACT

Title of dissertation: IMPROVING THE FOUNDATIONAL KNOWLEDGE OF DEPENDENCY IN HUMAN RELIABILITY ANALYSIS

Vincent Philip Paglioni, Doctor of Philosophy, 2023

Dissertation directed by: Associate Professor Katrina M. Groth
Department of Mechanical Engineering

Human reliability analysis (HRA) is the field tasked with understanding, modeling, and quantifying the human contribution to the reliability of complex engineering systems. Human machine teams (HMTs, the groups of operators and human-system interface technologies that together control a system) currently contribute to over 60% of industrial accidents and will continue to serve an important operational role in complex engineering systems. As a result, it is critical to develop robust methods for characterizing HMT performance and reliability. One of the factors limiting the technical basis of HRA is the treatment of dependency, how task performances and influencing factors are causally connected. Currently, HRA does not have a sound framework for conceptualizing, modeling, or quantifying dependency. The concept of dependency is poorly defined, the modeling is lacks a causal basis, and the quantification of dependency is unsupported by literature or data. This research closes these gaps in the foundations of HRA dependency by enforcing a rigorous, quantitative causal basis for the conceptualization and modeling of dependency.

First, this research addresses the definitional and conceptual foundations of HRA dependency to provide a consistent technical basis for the field. This work proposes a single, complete, and appropriate definition for the general concept of dependency; one that is rooted in causality. This research also provides definitions for dependency-related concepts from multiple fields including probability, statistics, and set theory. The definitional basis laid out by this work standardizes

the foundations of the field and promotes the ability to more easily translate between previously-disparate HRA methods.

Second, this work develops the causal structure of dependency in HRA. Whereas current methods for dependency modeling in HRA focus on correlational attributes, this method recognizes that causality, not correlation, is the driving mechanism of dependency. This work identifies six distinct relationship archetypes (idioms) that describe the general dependency relationships possible between HRA variables. Furthermore, this work creates the graphical structures that describe the idioms using Bayesian Networks (BNs) as the modeling architecture. The task/function-level idiom structures created in this work provide robust, traceable models of dependency relationships that can be used to both build HRA models and decompose full models into more understandable pieces.

Third, this work develops the methodology to build and quantify causal, formative dependency BN HRA models using the idiom structures and HRA data. Whereas many HRA methods rely on expert elicitation alone for assigning probabilities, this methodology quantifies the network directly from HRA data. The methodology developed in this work produces a full, causal, formative dependency scenario model without requiring expert elicitation of probabilities. This methodology is implemented to build and quantify a scenario model using real HRA data collected from operator crews working in a full-scope nuclear reactor simulator, which shows both that causal dependency can be modeled and quantified, and that the methodology is traceable and useful. Finally, this work develops a set of recommendations for the collection, storage, and use of HRA data, and for the implementation of this methodology within mature HRA frameworks. This dissertation will improve our knowledge of, and ability to model, dependency in human reliability.

IMPROVING THE FOUNDATIONAL KNOWLEDGE OF DEPENDENCY IN HUMAN RELIABILITY ANALYSIS

by

Vincent Philip Paglioni

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2023

Advisory Committee:
Associate Professor Katrina M. Groth, Chair/Advisor
Assistant Professor Michelle Bensi
Professor Mohammad Modarres
Assistant Professor Yunfei Zhao
Dr. Elizabeth Fleming, Special Member of the Graduate Faculty
Professor Gregory Baecher, Dean's Representative

Preface

Errare humanum est – to err is human – is Seneca's eloquent recognition of the human proclivity for error, and may be the first reference to an interest in *why* humans fail. The study of "human error" is an ancient venture, limited until relatively recently to the realm of philosophy. For the majority of human history, human error was simply the cost of being human, and had no real consequences external to the human making the error. Humans worked alone or manipulated simple technology with little feedback. However, the advent of complex engineering systems brought about a new context for humans: as one part of a symbiotic human-machine team and possessing the capability to cause tremendous adverse consequences as the result of an error. Accordingly, the study of "human error" evolved from pure philosophical inquiry into a *scientific* pursuit of not only *why*, but *how* and *how often* humans fail.

The study of human failure/reliability began in earnest with industries that still drive research in the field. Investment in human reliability seems to germinate in shorter intervals from the development of the technology according to the perceived catastrophic consequences of system failure. For instance, interest in human reliability analysis is first recorded roughly 100 years after the invention of the railroad, 50 years hence the introduction of the airplane and 20 years following the first commercialization of nuclear power.

The development of complex engineering systems has rendered the idea of "human error" something of a naïve mischaracterization regarding why systems fail. The term masks the fact that humans work both with and within systems to form a single "organism," and either part of the human-machine team can manifest a "human error." The recognition that *human-machine teams* – not humans alone – operate systems forms part of the context for this research, which follows in a long tradition of seeking explanatory theories for "human error." This research improves upon typical assessments of human reliability by investigating the human-machine team as a unit. Further, this research takes a causal perspective on human-machine team performance to explicate the web of causality that determines why, how, and how often human-machine teams fail.

I began the work that forms the substance of this dissertation in 2019, but my experience with nuclear energy systems and human reliability extends well before that point. I received my Bachelors' of Science in Nuclear and Radiological Engineering from the Georgia Institute of Technology in 2017. From 2017 to 2019, I worked as a Nuclear Test Engineer with the Department of Navy, wherein I received hands-on experience with the human side of complex engineering systems. This experience piqued my interest in how the human-machine team is understood and accounted for in engineering processes.

The research presented in this dissertation is the culmination of four years endeavoring to understand, model, and quantify human reliability. Understanding this esoteric aspect of complex systems results from significant introspection and an increased appreciation for the complex nature of how we as humans think, act, and interact with systems. The recognition that humans – like power plants – are complex creatures capable of failing in a myriad of ways is what makes this work so interesting, and so demanding. However, the purpose of this dissertation was, in part, to provide structure and traceability to the study of human reliability. Thus, if one uses the constructs and idioms created in this dissertation as a guide to structuring and evaluating human reliability, one will find that it becomes easier to trace the root cause(s) of errant behaviors. I believe that well-placed introspection is a powerful analytical, albeit inherently anecdotal or personal, tool for providing a familiar context to the ideas in this work. I urge readers, as they parse through this text, to imagine themselves in the example scenarios and consider how they might think, act, and respond if faced with similar circumstances. The processes and results discussed in this dissertation may further their understanding of their own actions and the web of causality that permeates them.

Acknowledgments

This research was supported in part by the U.S. Nuclear Regulatory Commission grant number 31310020M0002. Any opinions, conclusions, or recommendations expressed in this paper belong to the author alone and do not necessarily represent the views or opinions of the sponsoring agency. I would like to thank the A. James and Alice B. Clark Foundation for their support through the Clark Doctoral Fellows program at the University of Maryland.

This work would not have been possible without the support of so many people. I would specifically like to thank my Advisor, Dr. Katrina Groth, for her continuous support and contribution to my work, and her unfailing mentorship. My committee members, Dr. Michelle Bensi, Dr. Elizabeth Fleming, Dr. Gregory Baecher, Dr. Mohammad Modarres, and Dr. Yunfei Zhao, who collectively taught me about all of the different subjects necessary to performing this work, have been an enormous source of confidence and assistance. Additionally, thanks is owed to Dr. Michel Cukier, whose feedback and insights on tone strengthened the impact of this work, despite his inability to join the dissertation committee. The faculty at UMD have inspired me to continue my research and teaching as an academic.

My fiancée Jennifer Bates was and still is my biggest source of love and support. She gives me the confidence to perform this work and the courage to pursue my passions. Jenny, I love you and I thank you for putting up with me during this journey. It's been a difficult road to get to this point, and I could not have done it without you.

My family – my parents, my sisters, my brother, and my soon-to-be in laws – are my rocks, and I love them for their constant guidance and love, and for just being there for me when I needed them. Thank you to Mom, Dad, Sarah, Nicolette, Sofia, Sean, Scarlett, Anne, David, Julia, Jack, Jake, and Wheat for your constant support.

My friends in the Systems Risk and Reliability Analysis Lab, across the Center for Risk and Reliability, and around the University, have made this process bearable. You all taught me so much and made the campus a second home. I am very grateful to have gone through graduate school with all of you, and I look forward to the promises the future holds for us all.

Table of Contents

List of Tables

List of Figures

List of Abbreviations and Nomenclature

ASEP Accident Sequence Evaluation Program

ATHEANA A Technique for Human Error Analysis

BN Bayesian Network

CAP crew activity primitive

CES complex engineering system

CFM crew failure mode

CPT conditional probability table

CREAM Cognitive Reliability Error Analysis Method

DAG directed acyclic graph

EOC error of commission

HEART Human Error Assessment and Reduction Technique

HEP human error probability

HFE human factors engineering

HFE human failure event

HMT human machine team

HRA human reliability analysis

HuREX Human Reliability Data Extraction

IAEA International Atomic Energy Agency

IDAC Information, Decision and Action in Crew Context

IDHEAS Integrated Human Event Analysis System

IHA important human action

JHEDI Justification of Human Error Data Information

JRR *Journal of Risk and Reliability*

MCF major crew function

MCR main control room

NPP nuclear power plant

NRC Nuclear Regulatory Commission

PDF probability distribution function

PIF performance influencing factor

PRA probabilistic risk assessment

PSF performance shaping factor

SACADA Scenario Authoring, Characterization, and Debriefing Application

SPAR-H Standardized Plant Risk Analysis - Human Reliability Analysis

THERP Technique for Human Error Rate Prediction

Chapter 1: Introduction

Human reliability analysis (HRA), the study of human performance in a complex system, is a continuously-evolving field that is critical to the analysis of system safety in multiple regimes, but is particularly associated with nuclear power plant (NPP) operations. Essential to a correct understanding and modeling of human performance is the conceptualization and implementation of *dependency* in the HRA. Dependency (also referred to in statistics as *dependence*) in HRA is the process of determining how the probability of failure on one task changes due to failure of a previous task. However, this characterization of dependency, limited to tasks and without regard for causality, is incomplete for the HRA context. This research exposes foundational issues surrounding the conceptualization of dependency as a concept in the field at-large, as well as implementation-related issues in common HRA methodologies.

1.1 Motivation

Humans are a necessary facet in the design, operation, and maintenance of complex engineering systems (CESs), particularly nuclear power plants (NPPs). Moreover, industrial accidents often feature some sort of human component, with typical estimates ranging from 60-90% of events attributed to humans in some aspect of CES operation. For nuclear applications, roughly 60% of incidents reported to the International Atomic Energy Agency (IAEA) between 2015-2017 had some sort of human component, a number which has remained static for over a decade [1]. Despite the proliferation of Generation III+ and the rising popularity of Generation IV and other advanced reactor designs, the consequences of nuclear reactor accidents and the resultant regulations imposed on the industry make it unlikely that humans will be excised from the operation of NPPs.

Similarly, the increasing inclusion of automation in advanced NPP main control rooms (MCRs) is not removing human operators, but transitioning them to different roles, which may serve to increase the human contribution of incidents, due to the increasing reliability of automated systems and lower operator familiarity with the system operation [2]. Small Modular Reactors (SMRs), which offer the capability of consistent baseload power applicable to the emerging concept of "microgrids" will retain human operators [3]. Even micro-reactors and the so-called "set-and-forget" nuclear power plant designs (e.g., TerraPower, Oklo) will inevitably involve some human component in the design, operation and/or maintenance of the systems [4, 5]. Such designs do not represent the end of HRA, but an expansion in the scope and applications that must be considered [6]. Rather than fading in importance with the rise of automation, HRA will continue to play a prominent role in the safe and reliable design and operation of complex systems, and it is therefore imperative to ensure that the field stands on solid foundations.

This research is motivated by the recognition of multiple foundational issues that remain in HRA despite some 70 years of development. Many of these issues can be related to a general lack of conceptual consistency, particularly surrounding the critical aspects of task analysis (i.e., decomposing scenarios into observable blocks of human performance) and the treatment of *dependency* (i.e., the inter-influence of HRA variables on each other) [7]. It is well-known that the treatment of dependency is critical to the process of HRA, and therefore to the characterization and assurance of system safety and reliability; it is further recognized that the adjustment of the human error probability (HEP) for dependency has large-scale implications for the results of the HRA, and therefore the parent probabilistic risk assessment (PRA) [8–15]. However, dependency has never been rigorously defined in HRA, nor are methods for accounting for dependency based in causality or literature. The conceptualization and implementation of dependency in HRA has been heavily (nearly completely) driven by the initial dependency framework outlined in Technique for Human Error Rate Prediction (THERP) ([16]), which is the prototypical HRA method for nuclear power applications that has greatly influenced the development of HRA at-large. This dependency framework has survived to modern HRA implementations relatively unchanged, despite the lack of explicit causal or cognitive bases underlying the assignment of dependency "levels" and the arbitrary adjustment of probabilities based on the assigned dependency level [15, 16]. This research seeks to rectify these foundational gaps in HRA dependency treatment to enhance the development of quantitative HRA methods.

1.2 Technical Gaps in Human Reliability Analysis

The literature review (Chapter 2) revealed foundational technical gaps in the current conceptualization and treatment of HRA dependency, and the lack of current research addressing them. The current understanding of dependency is limited to a traditional view of direct dependency between HFEs (i.e., "error begets error"), and supported by the widespread implementation of THERP and its descendants, which typically provide vague guidance for assessing direct dependencies [12, 17, 18]. There are conceptual flaws now recognized in THERP dependency that have not been properly addressed, and have since been incorporated in many HRA methodologies that employ the same technical basis. This is coupled with the use of different terminology in various methods, and has resulted in confusion and division surrounding basic concepts related to HRA dependency [17, 18]. Implementing a more robust HRA dependency framework requires addressing three gaps in the conceptualization and application of dependency in HRA:

TG-1. The lack of a unified lexicographical and mathematical dictionary of fundamental terminology and concepts related to dependency in HRA (Chapter 3).

TG-2. The lack of an exhaustive and orthogonal set of dependency relationships capable of describing all relationships between HRA variables (Chapter 4).

TG-3. The lack of valid, causally-based mathematics to describe the effects of dependency relationships on the state probabilities of HRA variables (Chapters 5, 6).

The objectives of this dissertation (Section 1.3) were designed specifically to address these Technical Gaps, which are principally related to the conceptualization, modeling, and quantification of dependency. Closing these technical gaps will go a long way towards addressing a fundamental limitation in HRA which manifests as the technical gaps: the weak theoretical foundations for HRA dependency.

1.2.1 Fundamental Limitation: Weak Theoretical Foundations

Current HRA dependency treatments are largely based on the method originally laid out in THERP [16]. However, the THERP method does not have a rigorous theoretical basis from which the dependency levels, modifiers or equation(s) were drawn. As a result, THERP-derived dependency analyses have neither strong technical foundations nor have they been rigorously validated or verified [14]. Even HRA methods with dependency treatments sufficiently distinct from the original THERP (e.g., IDHEAS) suffer similar gaps in their technical foundations that must be addressed. The lack of a theoretical foundation for dependency modeling and assessment is a through-line in HRA as a field that manifests principally as three technical gaps.

1.2.2 Technical Gap 1

A review of the literature revealed that dependency as a concept was never rigorously defined in scholarly literature, guiding documents or methodologies for HRA. A *rigorous* definition entails explicating the causal, probabilistic and variable-agnostic nature of dependency (as will be discussed in Section 3.6.2). Definitions are available for dependency that are partially robust, in that they express some, but not all, of these aspects. The lack of a robust definition has hampered the development of an accurate dependency assessment methodology and has also allowed the proliferation of multiple, incongruous definitions for dependency.

Dependency is not robustly or consistently defined in common HRA method descriptions, which instead opt for vague conceptualizations that leave room for confusion and misinterpretation. Additionally, the concept of dependency is consistently limited to a relationship between HFEs or tasks. The quantitative impact of the change in state probability is not included in the definitions consistently.

Beyond the method descriptions, several industry standards and U.S. NRC regulatory reports

(NUREGs) have been issued in an attempt to address various gaps and synthesize a standard operating procedure for HRA. *However, these do not explicitly address dependency*. The IEEE HRA Standard, for instance, only defines dependency as the relationship between HFEs that "may" adjust the HEP [19]. The PRA Standard issued by the American Society of Mechanical Engineers (ASME) provides no explicit definition of dependency [20]. An early attempt to develop a method-agnostic HRA procedure provided no definition of dependency, presumably employing the THERP definition [21]. A subsequent set of NUREGs designed to establish good practices in HRA and evaluate HRA methods against the guidelines both did not define dependency [8, 22]. Two reports aimed specifically at addressing dependency in HRA define dependency by generally discussing the influence of an HFE on a subsequent HFE. One acknowledges the adjustment of the probability as a result of dependency ([23]), while the other limits the discussion to the general "influence" between HFEs [24]. Both frame the definitions with imprecise language rather than robustly asserting the qualitative and quantitative aspects of dependency [23, 24]. Relatively recent reports aimed at updating HRA in light of the digital I&C revolution and resurgence of Bayesian methods provide no definition of dependency [25, 26]. An NRC-sponsored training course in HRA provides one of the only examples of an explicitly probabilistic definition, but still limits dependency to acting between HFEs [27].

The closest that any HRA method, industry standard or NUREG comes to a robust definition of dependency is that it represents a relationship between HFEs or actions that *may* alter the HEP. This is not consistent with the statistical concept of dependence, which is applicable between any observable variables [28]. Nor is this definition consistent with the observed reality of HRA, where non-HFE variables (e.g., PIFs) can be dependent [29–31]. For instance, forcing independence between "Time Available" and "Operator Stress" is *a priori* inconsistent with reality, but HRA has typically treated PIFs as independent of each other — instead conceptualizing dependency as something that occurs only between tasks or human failure events (HFEs).

Even highly-relevant dependency literature does not put forward a concise definition of dependency, relying on the implicit assumption that it is a generalized "relationship" between actions

[32–35]. This is a foundational issue for HRA: dependency can have substantial consequences on the resulting HEP, and it is imperative that the field develops a robust, coherent and accurate definition of dependency that covers all aspects of what dependency means in an HRA context. The lack of a robust definition for dependency is emblematic of the more general lack of standardized terminology and mathematics in HRA literature and methodologies, particularly surrounding concepts related to task analysis and dependency.

1.2.3 Technical Gap 2

Another extension from the lack of a robust definition of dependency is the widespread disregard of causality and, to a lesser extent, cognition in HRA dependency. Whereas cognition and cognitive aspects of human performance have enjoyed a resurgence in recent HRA research ([36–38]), causality has hitherto been neglected as a driving factor for dependency. A coherent description of dependency must be cognizant of the cognitive aspects of human performance and combine them with the "machine" aspect of the situation in a causal framework. This combination is expressed partly in the use of the human-machine team (HMT) as the activity mechanism in HRA, recognizing that neither mechanistic/behaviorist nor cognition/cognitive aspects alone fully characterize human performance [30]. Moreover, coherent descriptions of dependency must recognize the importance of causality in the relationships.

The dependency framework set up by THERP, and now ubiquitous in HRA, can be described as a "positive dependency" or "error begets error" paradigm [39, 40]. This means that dependency is assessed *only* between failed tasks, with the assumption that a failure serves to increase the probability of subsequent failures (i.e., one error begets more). This approach emphasizes conservatism in the analysis, although that feature may be somewhat lost in implementation. Functionally, therefore, persisting under the "error begets error" paradigm masks the complex nature of interrelationships in HRA, which are neither as limited nor as linear as the current treatment suggests [15, 41].

Dependency as designed and implemented in HRA focuses on relationships between HFEs to the exclusion of relationships between any other HRA variables (Section 3.4.1). However, this belies the complicated reality that exists in HRA scenarios, wherein there is a complex web of variables and interrelationships that need to be accounted for in order to gain a realistic estimate of the HEP. In the SPAR-H method, for example, a HFE can be affected by any or all of the eight PIFs included with the method; other methods include similar PIF sets and recognize that a single HFE can be influenced by multiple PIFs [16, 31, 42].

The PIF-HFE connection is only one of the relationships between HRA variables that must be considered. There are also, for example, relationships between the failure modes (crew failure modes (CFMs)) and the HFEs, the major crew functions (MCFs) to their associated CFMs, and between the performance influencing factors (PIFs) and the other HRA variables. In current HRA, these relationships are largely neglected during dependency analysis, as are relationships between the PIFs themselves. The PIFs, in modeling the social, technical, environmental and personal factors weighing on human performance, describe a necessarily highly complex, dynamic and interrelated set of elements. However, PIFs have traditionally been assumed to be independent, in contrast with both HRA findings ([31] identified possible PIF interrelationships and their strengths, but neglects to consider these relationships in the method) and experience in reality. The reality of HRA is therefore a complicated web of variables and causal relationships, all of which must be captured in a robust dependency framework.

1.2.4 Technical Gap 3

The lack of rigorous theoretical foundations for dependency may be nowhere more evident than in the mathematics currently underlying most dependency quantification methods (Section 2.4.1). Currently, HFEs (or tasks) are judged against a set of correlational "dependency factors," and then assigned a dependency level, which in turn assigns a modifier to adjust the dependent HEP. However, these dependency levels and corresponding multipliers do not have a rigorous basis in experiment or literature. If the reason for including dependency in HRA is to impart conservatism and correct for uncertainty, this might be permissible on the basis of imparting sufficiently high conservatism (e.g., the rationale behind the HEART and JHEDI methods (Section 2.4.4)) [14, 15]. However, understanding dependency must be seen as a vehicle to greater causal insights and more accuracy for risk management, and not simply a way of imparting conservatism into abstract numbers. Furthermore, this process reduces the traceability and realism of the resulting HRA model, and hides the causal structure that is critical to understanding human-machine team reliability. HRA data is available and should be used to quantify dependency in HRA models. This research will investigate a methodology for quantifying dependency from data to replace the existing dependency levels and modifiers.

1.3 Objectives and Tasking

1.3.1 Research Objectives

To resolve the foundational gaps related to the treatment of dependency in HRA, this research aims to establish the definitional and mathematical foundations necessary for treating dependency in a comprehensive, causally-informed and analytically appropriate manner. This research also develops the structural framework to identify and model dependency relationships between HRA variables at multiple levels of abstraction, through the use of causal Bayesian Networks (BNs) models. Finally, this research develops the mathematical modeling framework required to quantitatively represent dependency in HRA.

The scope of the research can therefore be delineated into a series of objectives, the completion of which form their own achievements in developing the treatment of HRA dependency. Each objective is further broken down into a set of tasks that can be performed in semi-series to complete each objective. The objectives and their associated tasks are listed below. The objectives and workflow is visualized in Figure 1.1. Figure 1.2 depicts the relationships between the Research Objectives and Technical Gaps identified in Section 1.2.

- **Research Objective 1: Develop the lexicographical and mathematical foundations of dependency in HRA** through the standardization of definitions and mathematics for core dependency-relevant concepts. This creates a cohesive set of core definitions for foundational terms and concepts that HRA currently lacks. RO-1 explicitly addresses Technical Gap 1.

  - *Task 1.1*: Review literature for terminology and concepts in HRA related to dependency.
  - *Task 1.2*: Identify gaps in dependency-relevant definitions and concepts.
  - *Task 1.3*: Create unified definitional basis for HRA dependency (words and equations).
  - *Task 1.4*: Create causally-informed definition of *dependency* for HRA.

- **Research Objective 2: Develop a causal framework of HRA dependency idioms** that model relationships possible between HRA variables. The identification of causally-based dependency relationships moves HRA beyond the simplified notion of "error-begets-error" direct dependency to consider a comprehensive picture of the *causal* mechanisms at work in HRA that manifest as dependency relationships. The causal structures serve as building blocks from which a full causal picture of a scenario can be built, and serve as the foundations for the quantitative model of HRA dependency. RO-2 explicitly addresses Technical Gap 2.

  - *Task 2.1*: Critically analyze current HRA dependency modeling techniques.
  - *Task 2.2*: Review existing BN *idioms* and causal modeling.
  - *Task 2.3*: Analyze BN idioms for applicability to HRA dependency.
  - *Task 2.4*: Create HRA dependency idioms and corresponding BN structures.

- **Research Objective 3: Develop the mathematical framework for quantification** of the causal dependency relationships identified in Research Objective 2. The development of a mathematical framework rooted in causality replaces the currently arbitrary dependency mathematics with a scientifically-based, causally-informed methodology for computing dependent human error probabilities (HEPs). RO-3 explicitly addresses Technical Gap 3.

  - *Task 3.1*: Create causal BN model of HRA scenario case study, built from idioms identified in RO-2.
  - *Task 3.2*: Explore how to quantify causal BN model using SACADA data.
  - *Task 3.3*: Identify areas for continued study and/or improvement regarding HRA modeling and data collection.

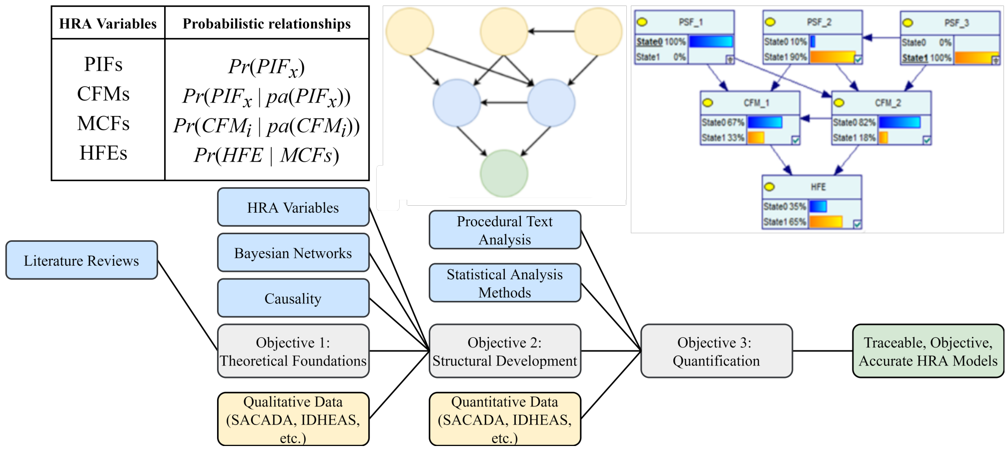


Figure 1.1: Dissertation objectives and workflow. Gray nodes are objectives, blue nodes refer to methods and techniques incorporated at each objective, and yellow nodes are data sources.

1.3.2 Research Objective 1 Tasks

Research objective 1 sets out the theoretical foundations of the research project. This objective develops the lexicographic and mathematical bases for dependency in HRA by providing standard definitions for common terms and concepts.

Task 1.1 reviews recent HRA literature to identify terminology and concepts in HRA that are in need of clarification. The literature search encompasses quantification variables and their definitions and task decomposition/analysis concepts, as well as mathematical and statistical concepts relevant to HRA dependency. Task 1.2 continues on this to identify gaps in the definitional and conceptual space from HRA literature. For instance, HRA variables that are multiply-defined (e.g., *PIFs* and *PSFs*) or poorly-defined (e.g., *HFE* and *dependency*), and statistical concepts that are commonly confused (e.g., *independence* and *mutual exclusivity*) require standard, robust definitions.

Task 1.3 towards this objective generates the definitions to resolve the gaps identified in Task 1.2. This task requires a thorough understanding of the *use cases* of each concept in HRA, which are available from the literature. For statistical concepts, which are non-controversial and generally have widely-endorsed definitions, Task 1.3 involves relaying the accepted definitions from statistics and mathematics to the HRA community.

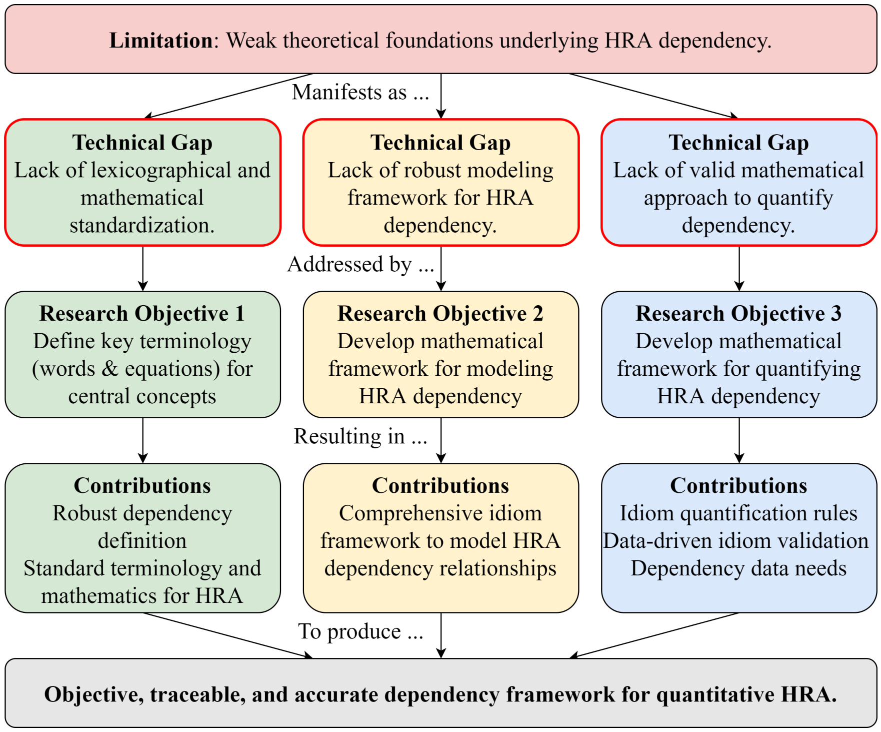


Figure 1.2: The relationship between Research Objectives, Technical Gaps and Expected Contributions.

The final task toward Objective 1 is the creation of a definition for *dependency* itself. As identified in the literature (Chapter 2), no HRA method or guiding documents robustly define dependency in a useful, scientifically sound manner. Task 1.4 uses the results of Task 1.1 to develop a robust definition of dependency (Section 3.6). This task is separated out from Task 1.3 to emphasize the greater level of creative effort required in developing a *new* definition over consolidating or finding definitions.

1.3.3 Research Objective 2 Tasks

Research objective 2 identifies the relationships possible between HRA variables, develops a set of fundamental relationship archetypes (idioms) built in BNs, and combines these into a full causal network to describe dependency in HRA.

Task 2.1 examines the current HRA dependency modeling techniques, principally THERP, SPAR-H and IDHEAS, to garner a greater understanding of both the possible HRA relationships and the "accepted" techniques in dependency modeling. Task 2.2 identifies the current uses of *idioms* and reasoning structures in building BNs for general-purpose models.

Task 2.3 uses the results of Task 2.2 and reviews the applicability of existing BN idioms and conditional reasoning to modeling HRA relationships. As shown in Chapter 4, not all of the identified idioms are applicable to HRA, nor do the idioms from ([43]) span the space of possible HRA relationships. Task 2.4 uses the results from Task 2.3 and creates a set of HRA-relevant BN idioms to describe dependency in HRA. The idioms created in Task 2.4 are fundamental logical structures that represent archetypical relationships between HRA variables.

1.3.4 Research Objective 3 Tasks

Research objective 3 develops the mathematics to quantify the causal dependency relationships from Objective 2. This objective moves HRA beyond the coincidental checklists and arbitrary levels and equations, and creates a sound model of the probabilistic changes engendered in the different HRA dependency idioms using a data-driven approach.

Task 3.1 begins by using the HRA dependency idioms to create a full BN model of an HRA scenario using SACADA qualitative and narrative data. Task 3.2 takes the network built in Task 3.1 and explores a methodology for quantifying the model (i.e., parameterizing the conditional probability tables in the BN) using the SACADA quantitative data.

Task 3.4 uses the process of building and quantifying the BN to identify HRA data requirements that, if incorporated, will facilitate more robust dependency quantification. Task 3.4 therefore leverages both Task 3.1 and Task 3.2 to recommend improvements to both HRA modeling and the collection, storage, and use of HRA data.

1.4 Impact

Humans will continue to play a role in the operation of NPPs and other complex engineering systems, regardless of the use of automated or passively-safe systems. In both the existing context and in the emerging paradigm of increasing disconnection between the human operators and "everyday" operation of the systems with which they are entrusted, it is important to have a robust understanding of the performance influencing factors (PIFs), error modes and mechanisms, and consequences of "human errors" for system safety. HRA as a field must grapple with the technical gaps undermining its foundations, and recognize the impacts these issues have on the ability to identify, contextualize, model and quantify human failure events (HFEs), particularly in high-consequence regimes such as nuclear power operations.

The research herein represents a significant shift in HRA as currently practiced, not only as related to the conceptualization and implementation of dependency in HRA models, but also addressing aspects of other foundational issues in HRA (Section 1.1). The completion of each individual objective detailed in Section 1.3 provides HRA with new capabilities that can be used to augment current practices in identifying, modeling and quantifying dependency. Research Objective 1 provides HRA practitioners with standardized language, definitions, and mathematical concepts to use in the performance of current HRA methods and development of new HRA techniques. Research Objective 2 creates a set of fundamental causal structures (the HRA dependency idioms) that describe the typology of dependency relationships possible between HRA constructs/variables. The HRA dependency idioms form the building blocks of objective, traceable, and accurate causal models for HRA, and conceptualize HRA dependency within an intuitive and causal framework. Finally, Research Objective 3 creates the comprehensive methodologies for constructing and quantifying data-driven causal HRA models using the HRA dependency idioms as the logical structure and eliminating the need for expert elicitation or dependency levels and equations in the quantification.

This research instills HRA with a causal basis from which to analyze and quantify dependency, as well as a definitional basis from which to build more cohesive and inter-consistent methodologies. This work also increases the role of causal quantitative modeling (via BNs) in HRA, improving the traceability and transparency of HRA models. To address the critical need for data in HRA, this research facilitates the standardization of HRA concepts to promote the collection and dissemination of critical data for validating the core assumptions and probabilities underlying HRA methods and models. Finally, this work provides recommendations for leveraging this work to improve HRA modeling and develop an improved HRA data lifecycle that will be more amenable to future modeling-focused research efforts. In total, therefore, this dissertation closes the three technical gaps related to HRA dependency and provides the capability to robustly conceptualize, model, and quantify HRA to gain a deeper causal understanding of why, how, and how often human error occurs in complex engineering systems. The practical impact of this research therefore extends well beyond the bounds of dependency to improve HRA as a general practice.

1.5 Dissertation Structure

The remainder of this dissertation is structured to provide a narrative review of the foundations of HRA, the foundational issues currently facing the field, and the research performed to fill the gaps related to dependency. Chapter 2 reviews the structure and scope of the background literature review, overviews the foundations of HRA, including popular methods and their inclusion of dependency. This chapter also details the limitations in the current treatment of dependency and the state of current research related to dependency in HRA. Finally, Chapter 2 reviews the current use of BNs in HRA as well as their application for modeling dependency.

Chapter 3 presents the results related to Objective 1, namely the lexicographical and mathematical dictionaries providing the basis for the new framework for dependency in HRA. This chapter presents a redefinition of the human failure event (HFE) to facilitate the top-down (functional-objective) modeling of human performance, and robustly defines *dependency* for HRA, contrasting it to both typical notions of dependency in HRA and the concept of statistical dependence. Chapter 4 develops the underlying causal structures for dependency, which are termed "HRA dependency idioms" and outlines how they can be incorporated into a full causal-BN model for HRA. Chapter 5 presents the methodology for building an HRA BN model using the HRA dependency idioms and HRA data, which was developed and refined with a case study scenario. Chapter 6 presents the methodology for quantifying HRA dependency (i.e., parameterizing the conditional probability tables in the BN) using the HRA dependency idioms and HRA data. Chapters 5 and 6 create coherent methodologies for construction and quantification of HRA BN models, built from the HRA dependency idioms, that were developed through the case study. Chapter 7 presents recommendations for leveraging the results of this dissertation to build more robust models and improve the collection, storage, and use of HRA data, that will facilitate future research in thsi area. Chapter 8 concludes with a review of the contributions, impact, and implications of this research.

Chapter 2: Background and Literature Review

This chapter discusses the initial literature review that was performed to identify dependency-related gaps in HRA. This chapter also provides a brief background on HRA as a field, including the typical objectives, major methods, and the state-of-the-practice and state-of-the-art dependency techniques in HRA. Finally, this chapter provides an overview of Bayesian Networks (BNs), including their current uses in HRA and possible applications to the modeling of HRA dependency.

This chapter is reproduced from portions of [18] and [44].

2.1 Introduction

Human reliability analysis is a field composed of several distinct methods for pursuing the same goals. HRA methods seek the theoretical goal of understanding why, how, and how often human error occurs, as well as the pragmatic goal of identifying the human error probability (HEP) that describes the probability of seeing an error on a given task or function. However, HRA is also beholden to both theoretical and practical technical gaps that limit its efficacy and applicability, both in general and related to dependency. This chapter provides a background of how and why HRA is performed, details the relevant literature regarding HRA dependency, and outlines the technical gaps that remain unresolved related to HRA dependency. Finally, this chapter provides an introduction to Bayesian Networks (BNs) as a modeling architecture for HRA. This chapter therefore shows that the extent technical gaps identified in Chapter 1 are not being addressed by either new HRA methods or current HRA dependency literature, and thus that there is a significant opportunity to improve the theoretical knowledge of dependency in HRA.

2.2 Literature Review Structure

The literature review for this research consisted of nominally two steps: a formalized, directed literature review search and an unstructured dynamic review of literature at-pace. This was done to gain a comprehensive understanding of the historical and current developments in HRA, particularly related to the conceptualization and implementation of dependency. The formalized literature review process gathered sources related to the historical context of HRA as applied in a nuclear context and the traditional use and development of the concept and implementation of dependency in HRA methodologies. At the same time, incorporating an unstructured search with regular literature updating kept the author abreast of the state-of-the-art research related to dependency in HRA.

The formal literature review process consisted of key-word searches in journal databases relevant to HRA. The literature search architecture implemented in this study consisted of three search passes interspersed with relevance filtering. The goal with this search was to thoroughly and exhaustively investigate the concept of dependency in HRA, and so the literature search was designed to collect the largest cross-section of literature reasonably obtainable, which could then be filtered for relevance according to the judgment of the authors. The resultant body of literature exhaustively covers the space of dependency research in HRA.

The first-pass search term was designed to capture the widest cross-section of literature that would be reasonably relevant to the subject of dependency in HRA, without producing an intractable number of irrelevant results. The first-pass search focused on journals related to HRA, reliability engineering and nuclear engineering, the field most closely associated with incorporating HRA. The first pass of the literature search was performed by searching the most relevant individual journals for articles where the title, abstract or author-designated keywords included the search term ((“human” OR “HRA”) AND (“dependency” OR “dependence”)), where “OR” and “AND” represent their respective Boolean operators. The journals included, and results obtained, are shown in Table 2.1. The first-pass search identified a total of 95 articles meeting the search

Table 2.1: Results from main journals during first- and second-pass literature search.

Journal	**1st-pass**	**Relevant**	**2nd-pass**
Reliability Engineering & System Safety	27	23	4
Safety Science	11	10	0
Journal of Risk & Reliability	40	4	0
Human Factors	1	0	0
Annals of Nuclear Energy	11	8	3
Nuclear Engineering & Design	3	3	1
Nuclear Engineering & Technology	2	2	0
Total	95	50	8

criteria, of which 45 were deemed irrelevant based on the title. Therefore, the first pass of the literature search yielded 50 relevant journal articles. It should be noted that the vast majority of irrelevant articles were discovered from the *Journal of Risk and Reliability* (JRR); this is due to idiosyncrasies in the Journal search engine which initially produced zero results for the initial search term. The JRR results obtained were found by searching ("dependence" OR "dependency") in the title, abstract and keywords, and as a result only a select few were relevant to HRA.

The second pass of the literature search consisted of searching the reference sections of each relevant first-pass article for relevant titles. This proved to be beneficial in identifying relevant and high-impact conference papers, textbooks and articles from journals not targeted in the first-pass search. The second-pass search identified an additional 24 sources, including additional papers from the first-pass journals, as shown in table 2.1. In addition, four relevant articles that had been identified in previous, less-formalized literature searches were added to the body of literature at this stage. A second filtering pass, conducted by reading the abstracts and performing in-text searches for the keywords 'dependence" and 'dependency" removed 16 irrelevant sources, resulting in a total of 58 relevant articles and two textbooks with relevant sections after the first and second search passes.

The unstructured portion of the literature review was performed by saving the Boolean search terms from the formal literature search process as Google Scholar search keys to give periodic updates when new literature relevant to the search terms was published. This facilitated the on-going collection of both conference and journal publications relevant to dependency in HRA, to augment the literature obtained in the structured literature review.

2.3 Human Reliability Analysis (HRA) Background

HRA is a process which focuses on the identification, modeling and quantification of "human error," particularly in the operation of CESs such as NPPs. HRA is typically used as an input and supporting method for a parent PRA, in which the HRA investigates the human-involved aspects of the scenario under investigation. HRA methodologies have undergone nearly continuous development in the decades since the introduction of THERP, the prototypical HRA method developed specifically for use in nuclear power [16]. HRA methodologies can be grouped roughly according to "generation," wherein the delineation is made by the introduction of new capabilities in each generation [39]. THERP and its related first-generation methods, namely Accident Sequence Evaluation Program (ASEP) and Standardized Plant Risk Analysis - Human Reliability Analysis (SPAR-H), largely ignored operator cognition. Second-generation methods such as Cognitive Reliability Error Analysis Method (CREAM) and A Technique for Human Error Analysis (ATHEANA) included more cognition, contextual factors and errors of commission (EOCs); the third generation of HRA methods, still under development, are working towards integrating aspects of dynamic modeling [39].

HRA is often connected to human factors engineering (HFE)[1] and/or ergonomics, which are nominally devoted to the *design* of human-interfacing machines and systems so that they match the capabilities of the human operators and facilitate human performance. HRA, on the other hand, is concerned fundamentally with analyzing the performance of the human *within* the system, and particularly on the consequences of "human error" on the system and scenario [16].

[1] Human Factors Engineering (HFE) should not be confused with Human Failure Event (HFE)

2.3.1 Objectives of HRA

The typical objective or output of an HRA is the calculation of the human error probability (HEP) (Section 3.4.1.7), which denotes the probability of human error on some critical task of importance to the parent PRA. However, the benefits of HRA are not fully realized without a robust underlying qualitative analysis to support the quantitative output of the HEP [31]. Nevertheless, the prevailing attitude within the PRA community is that HRA serves to provide the HEP as an input into the parent analysis. The full output of an HRA includes the HEP, but critically also entails developing a rich causal picture of the scenario, the important human actions (IHAs), relevant failure modes, contextual factors, dependencies and uncertainties [45].

HRA is fundamentally a qualitative and quantitative process that seeks to understand the impact of human interactions with the system (i.e., the NPP) on the system's reliability [31]. Any HRA is incomplete without *both* a robust understanding of the qualitative aspects of human performance in the system, and the quantitative output of the HEP for use in the parent PRA method. As a result, the HRA methods most commonly implemented in HRA all entail first a rich, qualitative process of information gathering and understanding of the scenario context, system operation, and human performance, after which the methods set out to quantify the probabilities of human error on given actions. The following sections overview the development and implementation of three major HRA methods with special emphasis on the conceptualization and implementation of dependency: THERP, SPAR-H and Integrated Human Event Analysis System (IDHEAS). These methods collectively span the entire "lifetime" of HRA as applied to nuclear power operations, with THERP representing the genesis, SPAR-H the most common operationalized HRA process in nuclear power (in the United States), and IDHEAS the most novel implementation of HRA [16, 31, 45].

2.3.2 Major Methods in HRA

2.3.2.1 Technique for Human Error Rate Prediction (THERP)

THERP was developed in the 1970s, and was implemented in nuclear power safety analysis with the publication of NUREG-1278, in the wake of the WASH-1400 report. NUREG-1278 served as the first handbook for HRA applied to nuclear power applications, and cast THERP as the genesis of HRA for nuclear power applications [16, 46]. Nearly every subsequent HRA methodology can trace some aspect of its construction to THERP [39]. THERP introduced what are now considered foundational concepts in HRA, particularly the use of performance influencing factors (PIFs)[2] for capturing the context of performance and accounting for dependencies between human error.

In the THERP framework, a "human error" is the standard analytical element, where a human error is defined as a human action that exceeds an acceptability limit, and therefore is an "out-of-tolerance" action [16]. Thus, THERP nominally addresses human performance at what is now considered the task or even sub-task level. To facilitate the HRA process, THERP provides a detailed process by which analysts are meant to investigate the situation, develop human event trees, and quantify the probability of task failure(s) based on nominal error rates, performance influencing factor (PIF) (also called PSF) states and ascribed dependency level. As part of this process, THERP provides multiple tables of standard human error probabilities (HEPs), which were summarized from the available literature and tabularized for use as baseline HEPs defined under nominal conditions. THERP also provides tables of HEP multipliers based on the assessed level of the observable PIFs (PSFs) *Stress, Experience*, and *Tagging Level* [16].

The methodology set down in the THERP handbook ([16]) is thoroughly-discussed, but difficult to implement due to the large number of quasi-subjective assessments that are required across multiple aspects of the analysis. THERP analysts are required to not only implement typical HRA elements, such as event trees, fault trees and possible human errors, but also identify the relevant

[2] also called performance shaping factors (PSFs)

PIFs, their states and associated HEP adjustments, and the level of dependency on prior tasks. The identification of relevant PIFs, states and corresponding HEP adjustments are performed by analysts with only high-level guidance from THERP, leading to significant effort required to develop an effective HRA, high subjectivity, and high inter-analyst variability.

2.3.2.2 Standardized Plant Risk Analysis - Human Reliability Analysis (SPAR-H)

Standardized Plant Risk Analysis - Human Reliability Analysis (SPAR-H) was developed in the 1990s and early 2000s to be a streamlined, operationalized version of the THERP methodology [31]. SPAR-H implements the salient aspects of THERP in a worksheet-based, user-friendly context aimed at increasing objectivity and reducing variability. To accomplish this, SPAR-H provides only two nominal HEP values, compared with THERP's fifteen tables tabulating HEPs for a variety of tasks, contexts and applications. Instead, SPAR-H designates a 'nominal" HEP ($nHEP$) for action tasks ($nHEP = 0.001$) and diagnosis tasks ($nHEP = 0.01$).

SPAR-H analysts are also provided a standard set of eight PIFs with possible states, consolidated from the over 70 pages of text recommending different PIFs and their possible states in THERP [16, 31]. SPAR-H also provides enhanced guidance for assessing the PIF states. Related to each PIF state in SPAR-H is a multiplier value used to adjust the $nHEP$ to create the 'basic" HEP ($bHEP$), which is essentially the conditional probability of human error given the recorded PIF states. Finally, SPAR-H retains the original THERP dependency framework coupled with a more concrete methodology for assessing the dependency level between two human actions.

One critical departure from the initial THERP conceptualization of HRA is the use of the human failure event (HFE) as the standard element of analysis, rather than the fuzzier 'human error." The term HFE suffers from its own definitional obscurity, as documented in Section 3.4.1.6, but represents a significant change from 'human error" in referring to a higher level of abstraction, i.e. 'function" rather than 'task" or 'sub-task." However, SPAR-H applies the THERP dependency framework proposed for inter-task relationships to inter-HFE relationships, leading to questions of the veracity of SPAR-H dependency calculations [37]

2.3.2.3 *Integrated Human Event Analysis System (IDHEAS)*

Integrated Human Event Analysis System (IDHEAS) was developed initially in response to a 2006 Nuclear Regulatory Commission (NRC) memorandum to investigate the current state of HRA methods in nuclear power. IDHEAS is meant to be a generalized methodology from which application- and scenario-specific HRA methods can be created [45]. The IDHEAS framework instills HRA with more recognition of operator cognition and errors of commission. The IDHEAS method includes three dependency contexts (consequential, cognitive, and resource-sharing) to build on the THERP and SPAR-H methodologies.

2.3.3 Technical Gaps in HRA

The review of literature related to existing HRA methods elucidated several technical gaps in the foundations of the field, including (but not limited to) the lack of consistent conceptualization and terminology for critical constructs in HRA, e.g., the PIFs, CFMs, MCFs, and HFE. Furthermore, the literature revealed significant gaps in the basis of dependency in particular, related principally to the definition of dependency, lack of causal modeling, and lack of data- or literature-driven mathematics for quantification, as discussed in Section 1.2. The technical gaps in HRA manifest as significant variations in the overall HRA result. As Figure 2.1 shows, the overall result of an HRA depends heavily on the method and team used. When applied to the exact same scenario, different HRA methods can vary by up to four orders of magnitude; different teams using the same HRA method can disagree by up to three orders of magnitude.

2.4 Dependency as Currently Implemented in HRA

Dependency analysis is a feature of most commonly-used HRA methods, although the inputs and process may differ across the methods. Briefly, dependency refers to the connection between HRA constructs; as in statistical dependence, there is an associated probabilistic change. The analysis of dependency is inextricably linked to the outputs of the HRA, and therefore to the

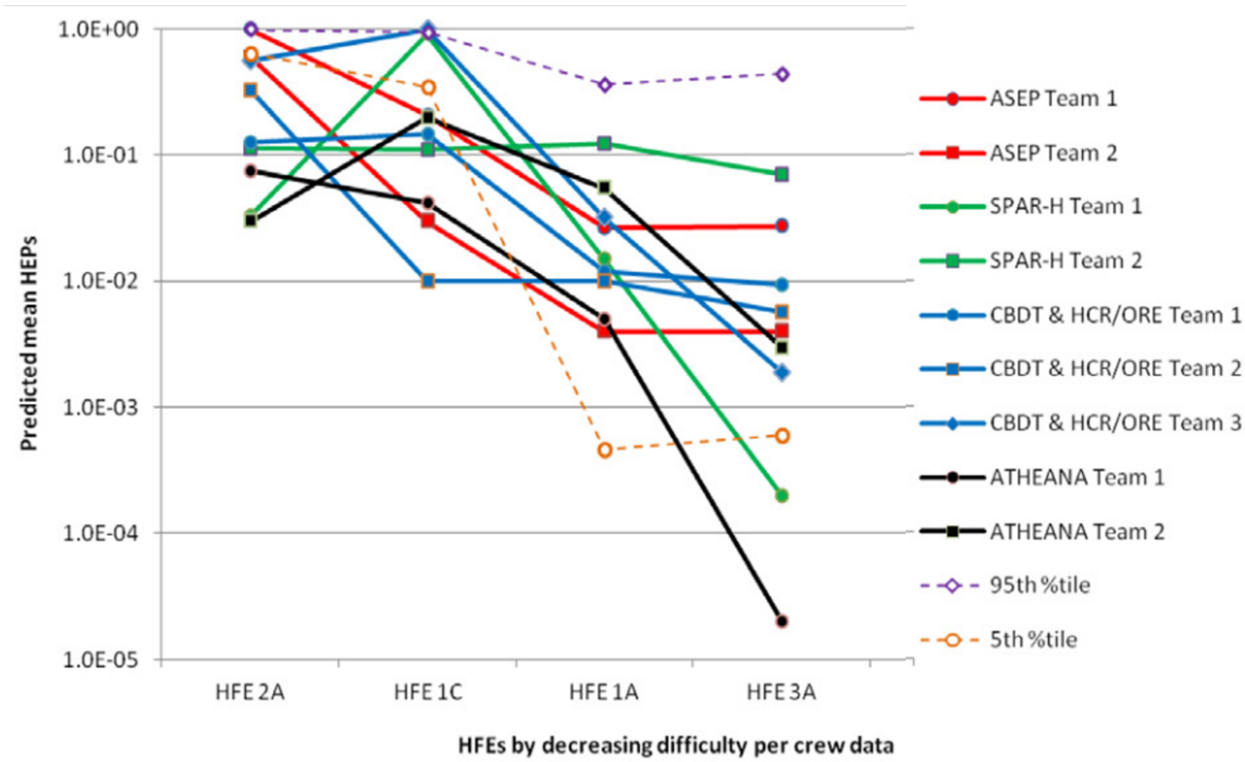


Figure 2.1: Cross-comparison of HRA methods and teams assessing the HEP on four HFEs. Adapted from [47].

outputs of the parent PRA. As a result, getting dependency ‘right" is of considerable importance in safety analysis. This section details the dependency methods in some commonly-used HRA methods, and evidences the outsized influence that THERP has had on the progression of HRA dependency, and HRA in general, since its publication.

2.4.1 THERP (1982)

THERP is arguably the first widely-implemented HRA method and, as a result, has enjoyed an outsized influence on the progression of HRA since its inception [16, 39]. THERP introduced the concept of *dependence* to describe the influence of human actions on each other, although a formal definition of dependency was not provided with the THERP method or subsequent documentation [16]. The assessment of dependency in THERP is based on a checklist of coincidental factors (e.g., temporal proximity, physical proximity, crew similarity) that determine the qualitative dependency

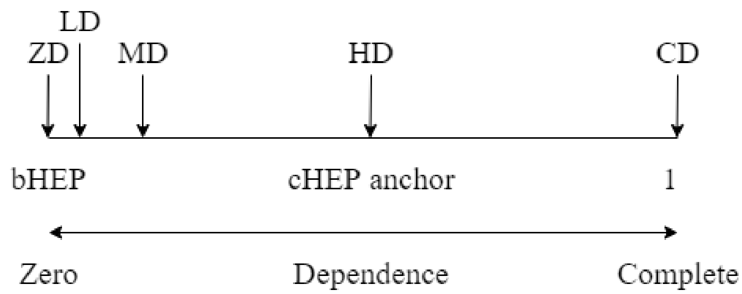


Figure 2.2: THERP dependency levels. Figure adapted from [15, 16]

level between two tasks (e.g., A and B) on a five-point scale from 'zero" to 'complete" (Figure 2.2). Each qualitative dependency level is associated with a 'multiplier" used to modify the baseline HEP (bHEP) of the 'dependent task" (i.e., B) to form the conditional HEP (cHEP). The cHEP of the dependent task is determined from the bHEP of the 'dependent task', and dependency multiplier k corresponding to the assigned dependency level. The cHEP of a dependent task B is given by Equation 2.1, wherein k takes on values $\{\infty, 19, 6, 1, 0\}$ corresponding to dependency levels zero (ZD), low (LD), medium (MD), high (HD), and complete (CD), respectively [16, 40].

$$cHEP(B) = Pr(B|A) = \frac{1 + k \cdot bHEP(B)}{k+1} = \frac{1 + k \cdot Pr(B)}{k+1} \tag{2.1}$$

The THERP method of dependency was developed for use between tasks, and provides a semi-traceable way for HRA analysts to generate conditional human error probabilities (cHEPs) based on factors assumed to associate a given task to a previous task. This same basic process, and the mathematics in Equation 2.1, form the basis of most dependency analysis methods in HRA.

2.4.2 SPAR-H (2004)

The SPAR-H method is perhaps the most widely-implemented THERP legacy HRA method, and in many ways is an operationalized version of the original THERP; for instance, the dependency method employed in SPAR-H is canonically identical to that in THERP. SPAR-H provides analysts with user-friendly worksheets to step through each step of the HRA process, including dependency assessment. The SPAR-H method incorporates the same factors for determining the dependency levels between two tasks as THERP (e.g., crew similarity, temporal proximity, physical proximity, additional cues) [31]. Notably, the SPAR-H method assesses dependency between HFEs rather than between tasks; the HFE became the standard unit of analysis in HRA following the publication of A Technique for Human Error Analysis (ATHEANA) [48].

Dependency assessment in SPAR-H is therefore a slightly more traceable and objective method than in THERP, but there are no substantitve changes to the *concept* of dependency provided in THERP. The SPAR-H method similarly does not provide a robust definition of dependency [31]. Accordingly, dependency in SPAR-H remains limited to inter-task relationships based on correlational connections between the tasks.

2.4.3 IDHEAS (2019)

The IDHEAS method published by the U.S. Nuclear Regulatory Commission (NRC) includes what may be the first substantive changes to dependency conceptualization and implementation in any HRA method [45]. IDHEAS describes three dependency *contexts* that determine the change to the scenario induced by an HFE by examining the cognitive structures underpinning the HFE connection to the system. IDHEAS defines consequential, resource-sharing and cognitive dependency contexts [45, 49]:

- *Consequential*: A preceding HFE directly affects aspects of subsequent objectives, such as definition/feasibility, Crew Failure Mode (CFM) applicability, critical tasks required and the time available for performance. For example, failing “Shut Isolation Valve A-1” may render the subsequent action “Open Valve A” impossible.
- *Resource-sharing*: Objectives share resources required for performance (e.g., system and human requirements). The occurrence of an HFE may result in fewer resources available for subsequent objectives. For example, if “Open Valve A” takes more time than anticipated, there may be inadequate time available for for the subsequent task.
- *Cognitive*: There is a “cognitive flow” between consecutive objectives (e.g., detection, understanding, responding, executing) that may be interrupted or changed by the occurrence of an HFE. For example, the crew may develop an incorrect mental model of the scenario following a failure.

Whereas THERP and SPAR-H focus on behaviorist or physical aspects of dependency, IDHEAS explicitly bases dependency assessment on cognitive factors in the scenario. IDHEAS uses the PIFs and contextual factors to drive dependencies as opposed to the checklists of dependency ('similarity') factors found in both THERP and SPAR-H. As a result, IDHEAS does not need to rely on arbitrary levels or equations for adjusting the HEP for dependency; instead, dependency affects the PIFs and CFMs weighing on a dependent HFE, and the updated HEP is computed based on the new states of these parent variables [45].

2.4.4 Other Methods

In addition to the methods described above, several other HRA methods have attempted to incorporate dependency into their analyses. These methods, including Accident Sequence Evaluation Program (ASEP), ATHEANA, Human Error Assessment and Reduction Technique (HEART) and Justification of Human Error Data Information (JHEDI), vary in their definition and implementation of dependency [50]. ASEP and ATHEANA provide definitions similar to THERP and SPAR-H, namely that dependency between tasks influences the probability of failure. The dependency analysis methods of both ASEP and ATHEANA draw on elements from THERP [48, 51]. JHEDI and HEART do not provide definitions of dependency. HEART foregoes the analysis of dependency, opting instead for conservative HEPs to implicitly capture the effects of dependency. JHEDI employs the THERP method in conjunction with limiting HEP values to induce conservatism [50].

2.4.5 Dependency Quantification Issues

The THERP dependency framework, continued in SPAR-H and other THERP legacy methods, not only features correlation-based (rather than causally-driven) dependency levels, but also a summative process of quantification [15]. In practice, this means that dependency is assessed after all other quantitative analyses – in SPAR-H, for example, the effects of the PIFs are computed before any dependency-induced changes in the probability are considered. However, as Boring notes, this means that assessing dependency produces anchoring values instead of unique probability adjustments – low values of $Pr(B)$ in Equation 2.1 (i.e., $Pr(B) \leq 0.05$) will result in a final cHEP equivalent to $\frac{1}{k+1}$ [15]. Because HEPs are typically low, this means that dependency quantification can override all previous analyses and render them unnecessary.

For a field that prizes analytical rigor, this result should raise concerns over the veracity of the mathematics used to adjust probabilities. Figure 2.3 depicts how this can affect the overall result of an HRA. If the HEPs assessed in Figure 2.1 were all adjusted for High Dependency in the THERP-derived framework, the order-of-magnitude variation disappears erroneously as the cHEPs are adjusted to the range of $\{0.5, 1.0\}$. This means that assessing a level of dependency between two tasks, regardless of the specific reasons for the relationship, will result in the same probability adjustment.

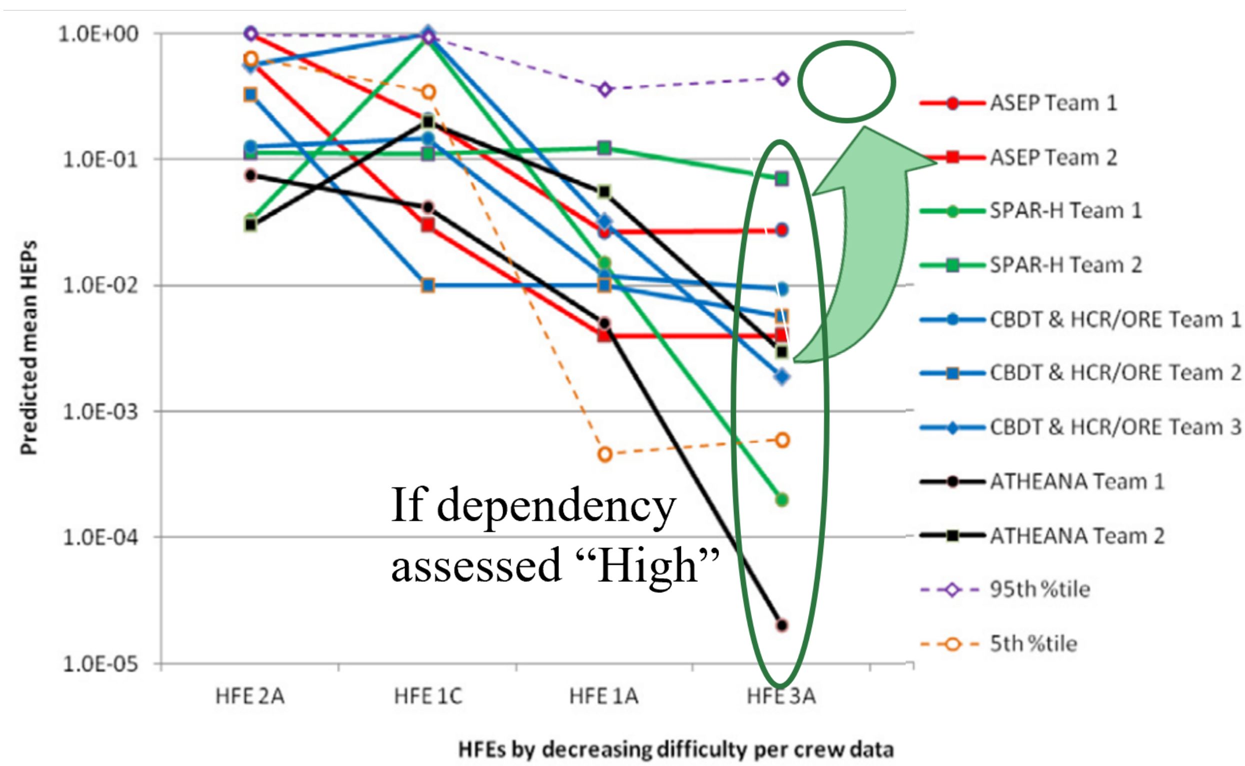


Figure 2.3: Effects of THERP dependency on final HEP values. Adapted from [47].

The idiosyncrasies in the dependency quantification are one manifestation of the technical gaps in the foundations of HRA dependency. Inconsistent conceptualization of dependency has led to conflicting understanding of what dependency means in HRA and how it should be assessed. Further, the lack of a causal modeling structure for HRA means that typical dependency assessments are correlation-based, relatively subjective and difficult to trace. Finally, the quantification of HRA dependency is not based in data or literature, and does not adjust probabilities in a sensible fashion.

2.5 Current Research Related to HRA Dependency

Despite a sizeable body of literature in dependency-relevant research in HRA, current work is focused largely on the pragmatic aspects of improving the subjective dependency assessments, rather than on correcting the technical basis of dependency. A much smaller subset of literature focuses on new techniques for *modeling* (graphically and/or mathematically) dependency relationships. Only a small subset of current literature in HRA addresses foundational questions relevant to dependency, such as what *dependency* should mean in the context of HRA and which HRA elements are subject to dependency relationships. As such, the flaws in the foundations of dependency for HRA are not being appropriately addressed in either the practical application of new methods or in cutting-edge research. This section reviews the current HRA dependency research corpus.

2.5.1 Application-based Research

The relevant articles identified in the literature search were overwhelmingly related to the application and/or improvement of existing dependency schema, and/or investigating best practices for the expert elicitation process. These sources investigated methods of "patching" the application of the existing (THERP) dependency framework with advanced elicitation/aggregation techniques, but do not address the weak technical foundations underlying their methods.

The current conceptualization of dependency, as shown in the literature review, remains heavily rooted in the guidelines developed for THERP. The dependency assessment methods of most HRA methods, then, require experts to assign correlation-based relationships between HFEs, and judge the strength of the relationship. Accordingly, dependency assessments are subject to high bias and uncertainty, and generally lack both traceability and repeatability [52, 53]. These limitations immediately impact currently-implemented HRA methods, and as a result a large majority of dependency-related research is devoted to developing methods to increase the traceability and objectivity of assessment methods.

2.5.1.1 Assessment Methods

Much of the research surrounding HRA dependency is focused on developing stop-gap measures to control and reduce biases and uncertainties induced by expert elicitation of dependencies. In current HRA dependency methods, the analysts must identify a relationship between two HFEs and specify the strength. Even with the checklists and tables provided with many HRA methods, there is an inherent degree of subjectivity in this process, that ultimately manifests as variability in the quantitative results. Several methods, reviewed briefly below, have been proposed to satisfy this need for immediate improvement.

Fuzzy logic is a set-theoretic methodology that allows for “degrees of truth” and approximate reasoning instead of rigid classifications [52–55]. Fuzzy logic dependency methods take the factors affecting dependency as inputs, and output an overall dependency level between two HFEs. Membership functions are used to transform the linguistic evaluations of the analysts into ranges of possible values (fuzzy sets), so that a single assessment of a factor state takes on an interval value instead of a point estimate. A set of “if-then” rules can be used to account for relationships between the factors [52]. A similar process is used for the dependency level, based on the assessed factor states, and the output is again a range of possible values, which can then be binned to a THERP dependency level. Fuzzy logic methods reduce subjectivity by allowing for range assessments rather than point-estimate assessments, and traceability is provided by the rule-based structure.

Evidence theory [56–58] methods are used similarly to fuzzy logic to reduce subjectivity in dependency assessments. Evidence theory can be used to aggregate weighted assessments from multiple analysts via linear or non-linear combination. The combined linguistic judgments are anchored to THERP dependency levels to produce a conditional HEP (cHEP) and confidence level. Evidence theory methods reduce subjectivity by providing a logical structure to aggregate assessments and using a computational model to develop the final cHEP. Traceability is provided in the method structure and documentation.

Analytic Hierarchy Process (AHP) and Analytic Network Process (ANP) are multi-criterion decision making methods that are used to aggregate assessments on multiple factors and across multiple analysts [35, 56, 59–62]. Like fuzzy logic and evidence theory methods, AHP and ANP are used to provide objectivity and traceability through a structured aggregation of factor assessments.

The improved methodologies for assessing the dependency level in HRA improve the objectivity and traceability of HRA dependency assessments. These methods aggregate subjective assessments to remove individual biases that may be present, but still map the final result back to a THERP-derived (or THERP-inspired) dependency level. By continuing to rely on unfounded dependency levels and equations, such methods do not functionally improve the technical basis of HRA dependency.

2.5.1.2 Modeling Methods

Beyond assessment methods, there are also efforts to improve the modeling of dependency relationships, mainly using Bayesian Networks (BNs). BNs are used to graphically visualize, and ultimately quantify, relationships between HRA constructs. The HRA constructs (e.g., PIFs, HFEs) are encoded as nodes in the BN, which are connected with directed edges representing the dependency relationships. BNs therefore provide a mechanism to model and quantify dependency relationships between *all* HRA variables, not just HFEs. BNs also eliminate the need for arbitary and unfounded dependency levels, relying instead on direct quantification of a cHEP based on empirical or elicited relationships. Groth and Mosleh ([29]) presented a methodology for deriving causal BNs for HRA that uses Factor Analysis (FA) to elicit inter-PIF dependency relationships. The emphasis on causality in this method is desirable for HRA in focusing on *real* relationships between PIFs and avoiding spurious connections. In their method, dependencies between HFEs and "error contexts" are driven by and through the PIFs, the traditional notions of "indirect" and "common-cause" dependencies [29]. Other works have used BNs to extend the SPAR-H methodology ([63]), to model cognitive causal pathways in HRA ([64]), and build causal models from event reports [65].

The use of BNs in HRA is growing significantly in interest, due in large part to their ability to inform both the qualitative (graph) and quantitative (CPTs) aspects of HRA. However, the most mature BN-HRA methods are still grappling with the explicit modeling of dependency. Further, some BN-based methods may enforce de-facto HFE independence by using separate BNs for each HFE [35, 39, 66–68]. However, this is not a feature of the implementation and not a limitation of BNs in general.

Advanced HRA modeling methods show great promise for improving the technical foundations of HRA dependency. BNs in particular, with their ability to visualize and quantify causal dependency relationships, represent a promising path forward for dependency modeling. As a result, BNs form the basis of the modeling structure in this dissertation.

2.5.2 Foundational Research on Dependency

Only a small number of articles were considered to be foundationally relevant to our work, meaning they were related to the theoretical basis of dependency. This small subset of HRA dependency literature is concerned with foundational questions surrounding dependency, including definition/notation clarity, the meaning of dependency, inter-PIF dependencies, and dynamic dependency. These works identify methodological differences driven by inconsistent notation, investigate PIF-driven indirect dependency relationships and inter-PIF relationships where the presence of a PIF changes the strength of another PIF's relationship to the HFE [35, 37, 40]. Because the THERP guidelines were developed for use in static HRA methods, others have worked to develop time-variant (dynamic) dependency relationships are also neglected [37, 41].

2.5.2.1 Definition, Notation and Parameterization Clarity

Some authors have recognized the foundational issues in dependency stemming from inconsistent notation and parameter definitions [33, 40]. Herberger and Boring noted the need for consistency and concision regarding probabilistic and statistical concepts related to HRA and expanded the THERP framework to include negative dependencies [40]. Čepin observed that the differences in dependency assessment methodologies of various HRA methods (including SPAR-H and IJS-HRA[3]) was the result of differing parameter definition and application [33]. De Ambroggi and Trucco found that the differences in the dependency modeling of several HRA methods (including CREAM[4], SPAR-H, and IDAC[5]) were driven largely by differences in the underlying conceptual models [35].

Čepin also presented a detailed analysis of IJS-HRA and SPAR-H viz each method's dependency assessment procedures [33]. Both methods incorporate the five dependency levels defined in THERP, but use different factors in the Decision Trees (DTs) used to assess the dependency level. IJS-HRA additionally differentiates between pre-initiator, initiator and post-initiator HFEs, where *initiator* refers to the immediate cause of an accident sequence [33]. The results of the case study showed that the dependency analyses differed between both methods and THERP itself; the differences were exacerbated when changing the number of discrete levels used to assess dependency. These results highlight the instability and subjectivity in THERP-derived dependency assessment methods, the impact that HRA inconsistencies have on the broader probabilistic safety assessment (PSA) results, and the necessity of standardizing parameter definitions and usage in HRA [33]. Čepin noted that the major differences between the methods, and the subjectivity of the results, were driven in large part by differences in parameter selection and definition, and the lack of detailed assessment guidelines; the problems associated with the lack of standard parameterizations was a common theme in the other foundationally relevant results [33].

The findings in these papers, further evidenced through a cross-sectional study of multiple

[3]Institute Jožef Stefan Human Reliability Analysis
[4]Cognitive Reliability Error Analysis Method
[5]Information, Decision and Action in Crew Context

HRA methods [47, 69, 70], confirm that there is an outstanding issue of clarity in fundamental aspects of HRA. Terms and concepts have been used inconsistently, and in some cases redefined completely, over the development of different HRA methods. Our previous work responded to these issues by developing lexicographical and mathematical dictionaries to standardize the use of critical concepts in HRA [44].

2.5.2.2 Meaning of Dependency

In addition to providing concise notation, Herberger and Boring investigated the THERP dependency framework and identified conflicts between probability theory and THERP outcomes [40]. Their work also proposed and quantified a negative counterpart to the THERP positive dependency framework, which describes the situation in which the probability of a subsequent task is negatively correlated to the outcome of the prior task. This extension of HRA dependency to include to negative dependency relationships fundamentally addresses the underlying conceptualization of dependency, i.e., what dependency *means* in HRA. This question is also addressed in ([37, 71]), which investigate dynamic dependencies driven by PIFs.

The first objective of this dissertation focused on the *meaning* of dependency, and identified the lack of a robust definition for dependency in any HRA literature [44]. The definition of dependency created herein (Section 3.6.3) is based in causality, explicitly probabilistic, and variable-agnostic, meaning that it comprehensively describes dependency for HRA.

2.5.2.3 PIF Dependency

Several authors have investigated dependency as applied to the PIFs, which have traditionally been treated as independent in HRA. Groth and Mosleh ([42]) developed a comprehensive, orthogonal taxonomy of PIFs for use in HRA and distinguished between the concepts of orthogonality and dependence. Groth emphasizes the distinction between *orthogonality* and *independence*; the PIFs are orthogonal, meaning they are uniquely defined, but will experience dependence relationships [42]. For example, “stress” and “time available” are *orthogonal*, meaning it is possible to uniquely assign observations to each PIF; However, these PIFs are not independent, as the time available for task performance will affect the level of stress during the task. Contrast this to "organizational factors" and "safety culture," for example, in which it is unclear whether an observed scenario aspect belongs to one or the other, meaning that these PIFs are not orthogonal.

Additional work with BNs explicitly discussed dependency relationships between PIFs [29, 30]. Several authors have investigated incorporating organizational factors into HRA as PIFs [41, 71, 72]. Boring recognized the importance of indirect dependency relationships driven through the PIFs, which are not considered in the THERP framework (which uses non-PIF factors to identify dependency). Boring also introduced the concept of "error spilling" and the potential of PIFs to exert contemporaneous impacts on each other [37]. Park, Boring and Kim expanded on this by developing dynamic dependency relationships between certain PIFs, driven by the biological processes underlying perception and experience [71]. De Ambroggi and Trucco proposed that PIFs can exhibit two types of dependency relationships on each other: between the PIF states, and between PIF effects [35]. They propose a model of PIF dependencies based on the ANP, however the main contribution to this dissertation is in expanding the concept of dependency to include relationships between PIF effects.

Boring, in the course of laying the foundations for dynamic HRA, critically noted that the THERP dependency framework largely excluded indirect dependency relationships driven by PIFs by focusing on direct dependency; this concept is adopted in this dissertation as *situational dependency* [37]. Some PIFs experience "lag" and "linger," where the effects of PIFs fluctuate dynamically and do not immediately switch on/off when "activated" (i.e., a stressor does not immediately result in maximum stress, but a temporal build up of stress to a maximum value). This induces time-dependent behavior to the PIF states and therefore to the relationships between PIFs; Park, Boring and Kim delved further into the dynamics of PIF effects by tying their impact to the biological processes governing perception (e.g., Cortisol absorption/secretion rates) [37, 71]. The quantitative study of dynamic PIF effects, using *stress* as a representative PIF is a promising effort to explore the possibilities of dynamic HRA modeling [71].

2.6 Bayesian Networks (BNs)

Bayesian Networks (BNs) are a class of probabilistic graphical models that incorporate a set of variables connected by directed edges to form a directed acyclic graph (DAG). The directed edges simulate relationships between the variables; causal BNs predicate the relationships on real causal mechanisms [73]. BNs have been heavily researched for applications in computer science and engineering analysis. This section reviews the background of BNs, how they are currently used in HRA and their potential applications in modeling HRA dependency.

2.6.1 Background on BNs

BNs are probabilistic modeling tools that use Bayesian statistical methods to update the probability of variables (expressed as nodes in the network) taking on certain states based on changes in related variables. BNs consist of two parts, the directed acyclic graphical structure showing the nodes and their relationships to each other via the directed edges, and conditional probability tables (CPTs) that tabulate the probability of observing the states of each node. Network nodes can be described as "root", "parent" or "child" depending on their position in the network with respect to other nodes [73, 74]:

- *Root node*: A root node is any node with no *incoming* directed edges (*A, B* in Figure 2.4).
- *Parent node*: A parent node is any node with at least one *outgoing* directed edge *(A, B, C, D* in Figure 2.4).
- *Child node*: A child node is any node with at least one *incoming* directed edge (*C, D, E* in Figure 2.4).

In Figure 2.4, the nodes[6] (*A, B, C, D, E*) represent instances of system constructs (e.g., performance influencing factors (PIFs)). Each node is defined as either continuous (representing, e.g., *Temperature, Probability* and other concepts that can assume a spectrum of values), where the state can take on any value in a defined interval, or discrete (representing, e.g., *Available Time, Stress*

[6] Here, "node" will be used when referring to the graphical representation; "variable" or "construct" is used to refer to the system element modeled by the node.

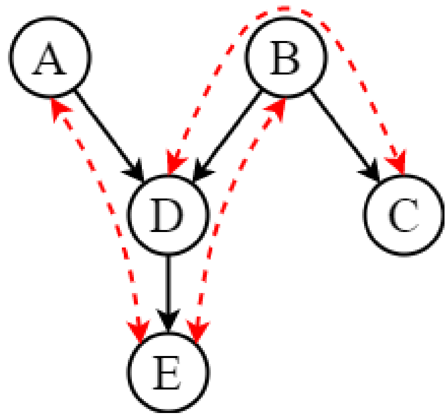


Figure 2.4: Example Bayesian Network graphical structure. Solid black lines are the directed edges that encode causal dependencies between the variables. Dashed red lines indicate available reasoning pathways when no evidence is provided to the network. Note that causality is unidirectional, whereas reasoning is not.

or other concepts that have individual states) with distinct possible states. Continuous nodes are defined with a probability distribution over the interval, and can take any value in the interval. Discrete nodes are assigned a number of distinct states, each assigned a discrete probability. This dissertation will focus on discrete nodes, which are common in HRA.

The directed edges in the graphical portion of the BN represent relationships between the nodes. A directed edge present between two nodes (e.g., $A \rightarrow D$ in Figure 2.4) indicates the presence of a causal relationship between the two variables represented by the nodes, with the arrow direction indicating the direction of causality (e.g., D is caused by A) [73]. The presence of a directed edge, and thus a causal relationship, between two nodes, changes the expression of state probability for the variables involved. Marginal probabilities (Section 3.5.2.1) are assigned to root nodes, while all non-root nodes are assessed with conditional probabilities (Section 3.5.2.3), conditioned on the state(s) of their respective parent node(s). The probability table of each root node therefore contains a number of cells commensurate with the number of defined states[7], while child node CPTs have $n \cdot \prod_i j_i$ cells, where n is the number of states for the child node and j_i is the number of states for each parent node i.

BNs not only offer graphical representation of causal relationships, but offer the power to reason about changes to system variable states using evidence garnered from a myriad of sources (e.g.,

[7]The probability distribution replaces the CPT for continuous nodes.

system sensors, known relationships, expert judgment). Using BNs allows analysts to leverage the causal relationships and the conditional independence between variables to facilitate modeling and quantification of the network.

2.6.1.1 Conditional Independence

Conditional independence, in the context of BNs describes the evidence propagation between three nodes on the basis of their relationship structure (i.e., the placement of directed arcs), which is described in Definition 2.1 [74]:

Definition 2.1. ***Conditional Independence***
Let $V = \{V_i\}, i \in \{1,...,n\}$ be a finite set of variables, and $P(\cdot)$ be the joint probability distribution over the variables V. Further, let X,Y,Z be any three subsets of variables in V. Then, sets X and Y are ***conditionally independent*** *given Z if:*

$$Pr(x \mid y,z) = Pr(x \mid z) \quad whenever \quad Pr(y,z) > 0$$

Conditional independence is denoted:

$$(X \perp\!\!\!\perp Y \mid Z)$$

Put another way, conditional independence between two variables means that evidence on one variable will *not* propagate to the other, because there is no pathway for propagation. As an example, consider a scenario wherein there are three tasks (T_1, T_2, T_3) performed in sequence, such that a simplified causal model of the tasks becomes a causal chain, as shown in Figure 2.5. Consider that there is evidence on task T_2. Then, regardless of new evidence on task T_1, the state probabilities of task T_3 will not change. This is due to the evidence on T_2 blocking the causal pathway between T_1 and T_3 [74]. The evidence on T_1 does not change our understanding of T_3, because the evidence already provided on T_2 has fully characterized the causal chain as regards T_3. In Figure 2.5, T_3 is said to be *d-separated* from T_1, meaning that all causal pathways (consecutive directed arcs) between T_1 and T_3 are blocked by the evidence on T_2 [74].

Conditional independence is also visible in Figure 2.4, in the diverging structure $D \leftarrow B \rightarrow C$. In this case, nodes C and D are *a priori* dependent on each other, due to the common cause in node B. However, once evidence is available on node B, nodes C and D will no longer affect each other, meaning that evidence added to either will not change the probability of the other. For example, consider the common-cause relationship $Cancer \leftarrow Smoking \rightarrow Yellow\ Teeth$, wherein smoking is a cause of both cancer and yellow teeth. With no knowledge of whether a patient is a smoker, the observations of cancer and yellowed teeth will be dependent (i.e., correlated). However, once it is known that a patient is a smoker, new evidence of cancer will not change the probability of yellow teeth, and vice-versa.

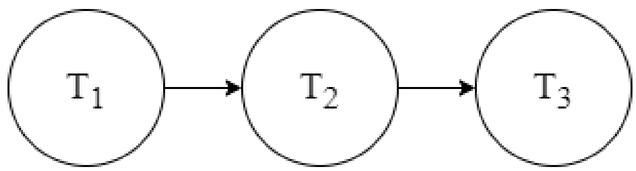


Figure 2.5: Simplified causal model of three sequential tasks.

2.6.1.2 Conditional Dependence

Closely related to the concept of conditional independence is the inverse concept of conditional *dependence*, which describes how causal pathways can be "unblocked" by evidence on a third-party node, and thus two otherwise-independent nodes are rendered *dependent* if information is available on a third node. Conditional dependence is best represented by a converging structure (e.g., $A \rightarrow D \leftarrow B$ in Figure 2.4); in this example, A and B are *a priori* independent, but become dependent once evidence is available on node D. For example, consider the common-effect relationship $Smoking \rightarrow Cancer \leftarrow Asbestos\ Exposure$. Both smoking and exposure to asbestos can cause cancer, but are *a priori* independent; evidence of either will not affect the other. However, if evidence of cancer is provided, then knowing the patient is a smoker will change the probability of their exposure to asbestos, and vice-versa.

2.6.1.3 d-Separation

d-separation describes how different logical structures on directed acyclic graphs can are affected by evidence, and therefore how evidence influences causal inference. d-separation is discussed formally via Definition 2.2 [74]:

Definition 2.2. ***d-Separation***

A causal pathway p is ***d-separated*** *by a set of nodes* $\{X\}$ *iff:*

- *p contains a causal chain* $i \rightarrow x \rightarrow j \quad s.t.\ x \in \{X\}$*, or*
- *p contains a causal fork (diverging structure)* $i \leftarrow x \rightarrow j \quad s.t.\ x \in \{X\}$*, or*
- *p contains a causal funnel (converging structure)* $i \rightarrow x \leftarrow j \quad s.t.$*:*
  $(x \notin \{X\}) \wedge \nexists\, y\, s.t. ((x \rightarrow \ldots \rightarrow y) \wedge y \in \{X\})$

The three logic proto-structures (causal chains, causal forks and causal funnels) described by d-separation form part of the foundation for the dependency logic structures (idioms) described in Chapter 4, and generally represent a spanning set of the possible logics relating any three BN nodes. Each of these can be seen in Figure 2.4:

- Causal chain: $A \rightarrow D \rightarrow E$; $B \rightarrow D \rightarrow E$
- Causal fork (diverging structure): $A \rightarrow D \leftarrow B$
- Causal funnel (converging structure): $D \leftarrow B \rightarrow C$

Expressing relationships in terms of d-separation allows for the quick recognition of the propagation of evidence via causal BNs; for simple networks by hand or simple algorithm (e.g., Bayes-Ball [75]), and for larger/more complex networks by computer program (e.g., GeNIe [76]). The implementation of dependency on Bayesian Networks is, therefore, intrinsically related to the concepts of conditional independence and d-separation. Current uses of BNs in HRA, however, are not focused explicitly on the modeling of dependency, but tend to translate existing HRA methodologies into the graphical formalism provided by Bayesian Networks.

2.6.2 BN Use in HRA

This section details the current uses of Bayesian Networks in HRAs, as well as the potential applications for BNs in conceptualizing and modeling HRA dependency.

2.6.2.1 Current uses of BNs in HRA

BNs have been widely incorporated in the performance of HRAs, including to expand the representation of relationships between HRA variables. BNs have been used to augment the performance of the SPAR-H method by converting the typical SPAR-H worksheet method into a graphical BN model that automatically updates state probabilities via Bayesian updating [63]. BNs have also been leveraged to expand the representation of crew failure modes (CFMs) by expanding the cognitive, causal pathways that manifest as various crew failure modes [64]. Other works have used BNs to explicate and model organizational factors, a currently vague, monolithic meta-factor that encompasses the work practices and culture of the organization and therefore likely has large influence on day-to-day work practices [77]. In the same vein, BNs have been explored for their ability to model the assumed interrelationships between performance influencing factors (PIFs) ([77, 78]). Bayesian methods, specifically Bayesian updating (Section 3.5.2.5) to adjust probabilities in light of new evidence, have been explored widely in HRA as well, particularly to develop better expert elicitation methods ([79]) and model the inter-crew variability in simulation performance ([80]). Despite the varied uses of BNs and Bayesian methods in HRA, there has not been an HRA-specific attempt to generalize and formalize the representation and quantification of dependency using BN structures. These are the subjects of Chapters 4 (the idioms), 5 (BN construction), and 6 (dependency quantification).

2.6.2.2 Application of BNs in Dependency

As shown in Sections 2.6.1.1 and 2.6.1.3, both evidence and the lack of evidence can enforce conditional independence relationships on BNs. Of particular importance to dependency in HRA are items (1) and (2) from Definition 2.2, the *causal chains* and *causal forks* (or diverging structures) logical protostructures. These are closely associated with colloquial notions of causality, referring to “cause-effect” and “common-cause” relationships, respectively [74]. In these relationships, the variables of interest (e.g., the termini nodes of the causal chain, and the effect nodes of the common-cause) are ground-state dependent, meaning the relationship exists when no evidence is available on the other nodes in the relationship. The final protostructure, the *causal funnel* or converging structure, refers to “common-effect” relationships; the variables of interest in common-effect relationships (i.e., the two (or more) cause nodes) are ground-state *independent*, meaning the dependency relationship exists *only* when evidence is available on the other node(s) in the relationship (i.e., the effect node(s)). For this reason, common-effect relationships are not typically counted among the pantheon of “dependency relationships” in HRA; note that none of the idioms described in Chapter 4 are common-effect relationships.

Bayesian Networks are therefore an excellent tool for modeling relationships in general, and this extends to modeling HRA dependency. BNs are central to this research because of their ability to portray and, through Bayesian updating, quantify the effects of dependency relationships. BN structures known as *idioms*, which are basic causal reasoning structures ([43]), were adapted to develop the HRA-specific idioms that model specific relationships present in HRA.

2.7 Conclusions

There are currently over 50 HRA methods available for the nuclear power context, although the most popular and most mature methods continue in the THERP tradition, particularly regarding the conceptualization, modeling, and quantification of dependency. For example, SPAR-H, perhaps the most widely-implemented HRA method, uses an identical method to THERP to account for dependency, and continues to address dependency only between tasks. However, recent developments in the IDHEAS method point to an interest in improving HRA dependency.

Regarding HRA dependency, there is a sizeable research corpus dedicated to its improvement. However, these works are focused mainly on improving the assessment and/or modeling of HRA dependency, while largely retaining the technical foundations developed in THERP. The majority of research relevant to HRA dependency is focused on developing and refining methods for aggregating expert-elicited dependency levels, but continue to rely on the THERP-derived dependency levels as the logical basis of the modeling and quantification. As a result, the impact of such research is hampered by the limitations inherent in relying on the existing, flawed conceptualization of dependency. It is clear that there are still technical gaps unaddressed, principally the lack of causality in the definition and in modeling, and the quantification of dependency. This dissertation was fostered by the realization that the main focus of HRA dependency research is not investigating the technical foundations of dependency.

There is a growing research thrust in the use of Bayesian networks for HRA modeling, which has the potential to drastically improve the treatment of dependency in HRA. Bayesian networks facilitate the visualization and quantification of causal dependency relationships, and accordingly form the basis of a large portion of this dissertation. The logical relationships defined in this chapter, i.e., conditional independence (Section 2.6.1.1), conditional dependence (Section 2.6.1.2), and *d*-separation (Section 2.6.1.3), will be revisited in Chapter 4 to set up the creation of the HRA dependency idioms.

This chapter showed that the body of research in HRA is not addressing the extant technical gaps related to the foundations of HRA dependency, but focusing on more pragmatic methods of correcting for theoretical shortcomings. Accordingly, there is a significant opportunity to close the gaps in HRA dependency related to the lack of a uniform definitional basis, lack of causal modeling, and lack of robust quantification methodologies. The remainder of this dissertation details the work performed to comprehensively address these technical gaps. The uniform definitional basis for HRA dependency is created in Chapter 3. The HRA dependency idioms created in Chapter 4 provide the causal modeling framework to describe dependency between HRA constructs, which is then incorporated into comprehensive methodologies for the construction and quantification of HRA BN models in Chapters 5 and 6, respectively.

Chapter 3: Definitional Basis for HRA Dependency

This chapter addresses Research Objective 1, the development of lexicographical and mathematical dictionaries to standardize concepts, terminology and mathematics relevant to HRA dependency. Therein, this chapter develops the unified definitional basis for HRA and outlines the requirements for a robust definition of dependency based on evaluations of existing dependency definitions. Finally, this chapter provides a definition of HRA dependency that is based in causality, explicitly probabilistic and variable-agnostic.

This chapter is reproduced from portions of [7], [18] and [44].

3.1 Introduction

One of the significant technical gaps in HRA, identified in Chapter 2, is the lack of standardized, comprehensive definitions for multiple terms and concepts foundational to HRA dependency. This lack of consistency hampers the development of HRA as a field, as methods and models built with different constructs are difficult to cross-reference and validate. Thus, there is a critical need for a uniform definitional basis in HRA that provides standard conceptualizations and definitions for the modeling constructs, language and mathematical relationships relating to HRA dependency. One of the critical definitions that is lacking in HRA is a robust definition for the concept of dependency itself, which is a poorly-defined and inconsistently used construct in HRA methods and literature.

Accordingly, this chapter identifies the lexicographical and mathematical concepts that require clarification and standardization to support HRA dependency, and defines them for the HRA context. Further, this chapter identifies the requirements of a robust definition of HRA dependency and provides a definition that meets these requirements and forms the first comprehensive definition of the concept for HRA. The concepts defined in this chapter provide a uniform definitional basis for HRA that will strengthen the technical foundations of the field and help to alleviate the variations in methods and results engendered by previous, siloed approaches to HRA development. Finally, the robust definition of HRA dependency created in this chapter allows for a uniform, causally-informed and technically sound conceptualization of dependency for HRA, which facilitates modeling dependency to gain deeper causal insights into why, how, and how often human error occurs in complex engineering systems.

3.2 Approach: Necessity of Clarifying Foundational Concepts

The literature review (Chapter 2) revealed that a lack of technical standardization is a factor underlying the persistence of many technical gaps in HRA. Therefore, standardizing the terminology and mathematics related to HRA dependency furthers the standardization of HRA in general. Many of the terms and concepts presented in this chapter existed, at least in principal, within the body of HRA work. In such instances, this chapter focuses on unifying those definitions drawing on their similarities and accepted concepts in HRA, rather than developing new definitions out of whole cloth.

The persistence of mutually-incongruous terminology has led to the gradual departure of HRA methods, manifested as significant variation between methods when applied to the same problem (even by the same practitioners), as shown in Figure 2.1 [47, 69, 70]. The history of HRA is rife with iteratively developing new, incrementally different methods to fix issues in the methods as they are identified; the result after 60 years of this iterative, siloed approach is more than two dozen HRA methods developed specifically for nuclear power applications.

The plethora of HRA methods in itself is one issue, but the far deeper problem is revealed when attempting to compare results across methodologies. The wide variation of terminology and processes employed in different HRA methods means methods yield very different results, a fact that has not significantly changed in 40 years [47, 69, 70, 81]. Moreover, the methods appear largely to be mutually-incompatible, meaning that they incorporate different assumptions, base probabilities and levels of detail that render them almost incapable of being compared.

The collection and analysis of data is a central component in developing any HRA method, particularly when trying to populate databases or validate the method. The mutual incompatibility of the majority of HRA methods means that available data sources also tend to be incompatible. Plainly, data collected for use with one HRA method is often not applicable to any other method unless great care is taken to translate the data. However, because the differences in methods are structural, the translation of the data is not a straightforward task, even if it is strictly possible. Issues apparent in HRA data are addressed further in Chapters 6 and 7.

One step towards solving these interrelated issues is, of course, standardization of the practice of HRA. This begins with standardizing the foundational concepts in HRA. Developing an HRA *lingua franca* will facilitate more effective translation between methods and data sources, but more importantly provides the theoretical basis required for the bulk of this research in developing objective, traceable and accurate HRA models.

3.3 Methodology: HRA Concepts Requiring Standardization

Currently, the most widely-implemented HRA methodologies remain significantly tied to theoretical and/or practical elements from THERP that have changed little in the intervening years [39]. Additionally, because of shifts in guidance and attitudes external to any HRA methodology, these methods can include inconsistent processes and produce conflicting results, despite their common roots [47, 69, 70]. Both the U.S. NRC and the IEEE have issued standards intended to regulate the performance of HRA and provide guidance to practitioners, but neither has succeeded in developing a unified foundational knowledge base [8, 19]. Two recent HRA methods, Phoenix and IDHEAS (Integrated Human Event Analysis System), have been developed to meet the demand for greater consistency, with an emphasis on cognition in decision-making and causal modeling of performance and dependency [30, 45, 82]. However, until these methods are settled in the HRA space, they are two additional sources of conflicting terminology and mathematics.

The continued presence of conflicting terminology and mathematics has plagued HRA since its inception, and has only grown in the ensuing years with the continuous development of new, incrementally-improved methods. This has resulted in several fundamental aspects of HRA being poorly defined, misused or misconstrued. This is despite the understanding that, similar to the use of dependency, the definition and implementation of HRA elements has a large effect on the overall results and meaning of the HRA and parent PRA analyses. To this end, the following sections review the current state of understanding and associated technical gaps for three "classes" of HRA concepts: HFE/HEP, independence and mutual exclusivity, and additional HRA variables/constructs.

3.3.1 Human Failure Event and Human Error Probability

Bridging the gap between the misconceptions surrounding dependency and task decomposition are the ideas of "Human Failure Event" (HFE) and "Human Error Probability" (HEP). The HFE is generally referred to as the highest level of task decomposition, and traditionally the only variable to which dependency is applied in HRA [31, 82]. HFE is one of the most commonly-used terms in HRA, and therefore it is essential that there is at least consensus regarding its use, if not a universal definition applicable regardless of methodology employed.

The HFE, as a concept, was developed with ATHEANA ([48]), while the earlier THERRP focused on the task-level concept of "human error." However, THERP does refer to the idea of task decomposition and estimating the failure probability of "larger units of behavior corresponding to entire tasks or groups of tasks," which may correspond to what is now considered an HFE [16]. SPAR-H defines the HFE as a basic event representing some human-caused failure of a component, system or function [16, 31]. SPAR-H also considers the HFE as the culmination of some number of task errors and Unsafe Actions. At first glance, the two definitions seem compatible, however a strict reading of THERP would allow an HFE to be assigned as the failed state of a single task, which SPAR-H does not seem to allow. An HFE in the SPAR-H sense could represent a much broader concept than the THERP analogue, a realization which has already led to concerns over the use of THERP-derived dependency computations in SPAR-H [37, 83].

Phoenix treats HFEs in a similar manner to SPAR-H, developing them top-down and referring to the parent PRA method for an appropriate level of decomposition [82]. However, this means that there is no explicit guidance in Phoenix regarding HFE construction, which leaves the method open to significant variation depending on the resolution of the PRA method with which it is coupled. Variation in HFE definition is not unique to Phoenix, many HRA methods rely on the accompanying PRA method to define the appropriate level of the HFE [83]. IDHEAS also considers HFEs to be the combination of several tasks, but again does not provide a concise definition for what should be considered an HFE. IDHEAS relates HFE to the concept of "Important Human Action," which is similar to the idea of Major Crew Functions (MCF) in the HMT-centered HRA framework, which is discussed in the next section and in [18, 30]. The definition and usage of HFE varies between methods, but even the same method employed by different teams often produces significant variation depending on how each team defines HFE; Recent cross-sectional HRA studies required pre-defined HFEs to facilitate even basic comparisons between HRA analysts and methods [47, 69, 70].

The general conceptualization of an HFE as the composition of multiple tasks is consistent with the HMT-centered HRA paradigm and the current thrust of HRA [18, 30]. The HMT context of HRA views the HFE as a process rather than a single event, which emphasizes the synthetic nature of the HFE and recognizes the reality that serious failure events are often the culmination of several smaller failures. This recognition of the HFE as a failed state or failure process, as opposed to an entity in its own right, is one consistent point between THERP, SPAR-H and Phoenix. IDHEAS does not distinguish HFE as exclusively the failed state of an action (or series of actions), but treats it as its own concept. The HFE should be viewed as the result of the failure process, rather than a distinct, performable construct that can be successful or failed. This may be a semantic difference, but is emblematic of the larger divisions and misconceptualizations driven by a lack of consistent terminology in HRA.

Human Error Probability (HEP) is inextricably linked with HFE, but the two are not interchangeable. The HEP, as the name implies, is the quantification of the probability of the associated HFE, while the HFE represents the process of failure itself. Simply, an HFE is the failure, while the HEP is the probability of the failure. Figure 3.1 visualizes the relationship between HFE and HEP, specifically the empirical HEP which is the fraction of failures compared to number of total attempts.

Beyond HFE and HEP, other fundamental concepts are often misused or misnamed, leading to more confusion and less cohesion in the field. Misnamed terms by themselves are an inconvenience, but a pattern or trend of continually renaming similar or identical concepts causes divergence over time, resulting (long-term) in the mutual incompatibility observed between current HRA methods. Misnamed terms are a long-term concern, but misused concepts are an immediate problem, indicating a lack of understanding around fundamental ideas in HRA. Both instances can and should be addressed by ongoing efforts to unify the field and introduce quantitative HRA [18].

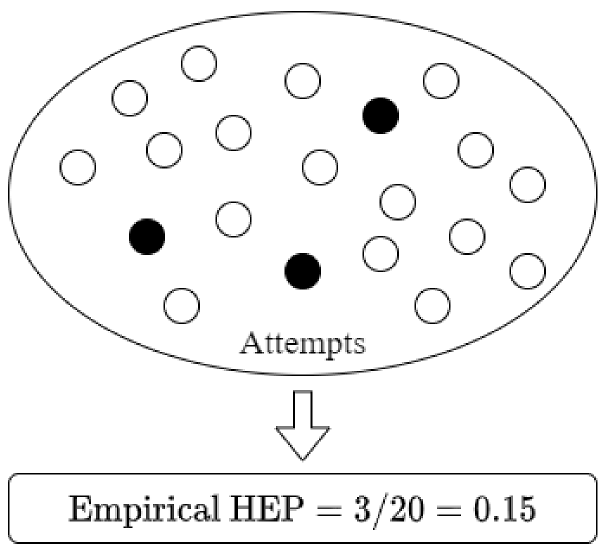


Figure 3.1: Visualization of the differences between HFE (filled circles) and HEP. HFE refers to the failed objectives as events, while HEP refers to the probability of failing an objective.

3.3.2 Independence and Mutual Exclusivity

Tightly coupled with the concept of dependency are the set theoretic concepts of mutual exclusivity and independence. Most HRA methodologies have consistent formulations of independence as distinct from mutual exclusivity. However, there is some evidence of confusion when it comes to the field at large, particularly concerning the visualization of each concept [40]. It is easy to confuse these concepts when trying to visualize them with a Venn diagram, because this is an inappropriate visual for dependency. Venn diagrams cannot distinguish dependent variables from those variables that can occur at the same time, which is not an indication of dependence. Instead, a better visual for dependency is an influence diagram or other causal model, that tracks the relationships between variables.

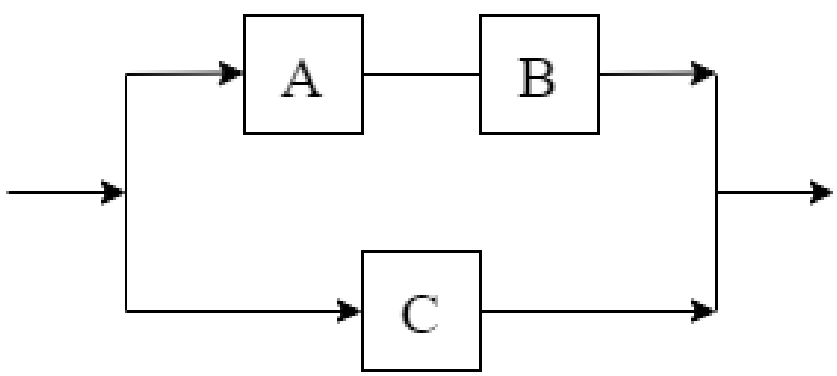


Figure 3.2: Reliability Block Diagram for hypothetical three valve system. Valves *A* and *B* are independent, while valves *B* and *C* are dependent.

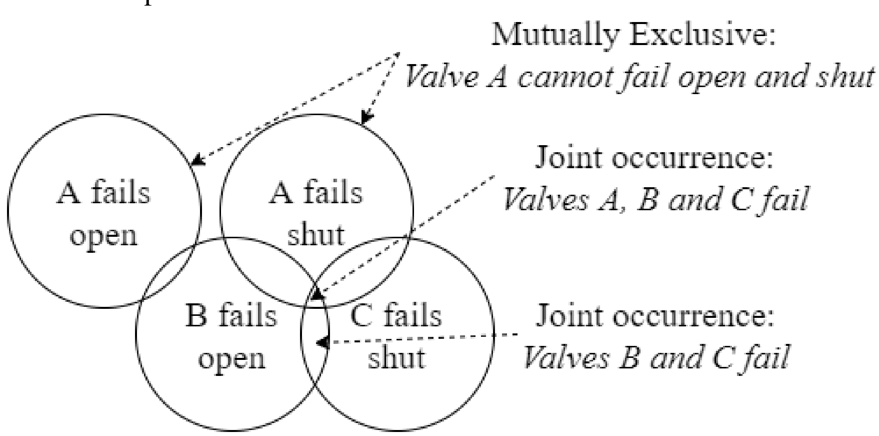


Figure 3.3: Venn diagram describing the system from Figure 3.2. Note that the Venn diagram cannot distinguish independent events (e.g., A and B fail) from dependent events (e.g., B and C fail).

The distinction between independence and mutual exclusivity is important for framing the relationship between two HRA variables both conceptually and quantitatively. Conceptually, independence and mutual exclusivity exist at different points on the same continuum of dependency (e.g., Figure 2.2), where mutual exclusivity is a type of complete dependence, but independence is zero-dependence. Variables that are independent can still exist simultaneously, while mutually exclusive variables cannot coexist and are necessarily dependent. For example, consider the system of three valves *A, B* and *C*, the Reliability Block Diagram of which shown in Figure 3.2, where valves *A* and *B* are independent, but valves *B* and *C* are dependent due to a shared control system. The valves may each fail open or shut. If a given valve fails shut, it cannot also fail open. The two outcomes for a single valve are not independent: the probability that the valve fails open given that it failed shut is zero, which is not equal the marginal probability of failing open, except in the trivial case in which the marginal probability is zero. The two outcomes are, accordingly, mutually exclusive. Valves *A* and *B* are independent but can fail simultaneously; they are not mutually exclusive. Valves *B* and *C* are dependent and can fail simultaneously. One Venn diagram of this system is shown in Figure 3.3, which cannot discriminate between the independent failures of *A* and *B*, and the dependent failures of *B* and *C*.

3.3.3 HRA Constructs/Variables

There are several elements of any situation that must be modeled and quantified in an HRA context. These are namely the HFEs, the subdivisions of HFEs (MCFs, CFMs, CAPs) and the performance influencing factors (PIFs), which will be referred to collectively as "HRA constructs" or "HRA variables" [30]. As Chapter 4 will show, each of these constructs can be subject to and can exhibit dependency relationships [17, 18, 30]. The HRA constructs in the HMT-centered HRA paradigm were made explicit in [30] but are still susceptible to confusion. In particular, the HRA Variable "MCF," Major Crew Function, is sometimes confused with, or replaced by, the similar concept of macrocognitive function, which refers to the high-level cognitive process archetypes used in problem-solving [2]. As Section 3.4.1.5 will show, the MCFs differ from macrocognitive functions in that they refer to the high-level, system-relevant actions required by the crew, not cognitive processes. MCFs and CAPs (Crew Activity Primitives) are built on macrocognitive functions, but are system-specific instances of a macrocognitive function applied to a particular scenario [30].

Performance Influencing Factors (PIFs) are the HRA construct most often referred to with conflicting language. In various circumstances and methods, these are referred to as Situational Factors, Performance Shaping Factors, and Error Forcing Conditions, among others [16, 31, 84, 85]. The differences between these terms mirrors a larger discontinuity in the definition and assignment of PIFs, which are broadly defined as the causal factors weighing on HMT performance and affecting the definition of MCFs and CAPs [30]. A single term for these variables is useful to coalesce the understanding around causal factors affecting HMT performance, and implement the universal taxonomy of PIFs presented in [42].

The differences in conceptualization and terminology between HRA methods has only exacerbated functional discontinuities and resulted in a field where different methods applied to the same problem can yield vastly different results (Figure 2.1). In addition, there is a serious need for review on the fundamental mathematics underlying HRA and specifically dependency, which is growing to be a serious concern for the field, as Quantitative HRA will rely heavily on a robust conceptualization of dependency which does not currently exist [18, 83]. This dissertation therefore created the following lexicographical and mathematical dictionaries to set a solid foundation for the push towards Quantitative HRA.

3.4 Result: Lexicographical Dictionary of Foundational Concepts

This section provides unified definitions for important concepts in HRA, focusing on those concepts relevant to HRA dependency.

3.4.1 HRA Constructs/Variables

HRA construct (or HRA variable) is used to refer collectively to the elements included in an HRA model in our analytical paradigm (e.g., PIF, CAP, CFM, MCF, HFE). These elements represent the contextual factors, functions, and failure modes critical to developing a causal model of any HRA scenario. Dependency is possible between any HRA constructs, so there is a need for a collective noun to make reference easier. The core HRA constructs defined here are summarized and visualized in Table 3.1.

3.4.1.1 Human machine team (HMT)

HMTs are groups of human operators and systems found in complex systems. The HMT works symbiotically to perform functions and accomplish objectives necessary to maintain system safety and/or operation. The roles of both the interface systems and the human operators are important in the context of human failure events, as the failure of either HMT component to perform its requisite function can result in an HFE [13, 30].

3.4.1.2 Performance influencing factor (PIF)

PIFs are the multi-dimensional characteristics (e.g., task, personal, crew, organization, or environment) that describe the *context* in which the HMT operates. PIFs, which are also called performance shaping factors or PSFs, have been used HRA methods since THERP to adjust the HEP or probability of failure [30, 42, 67, 86]. However, many methods rely on bespoke PIF taxonomies of varying cardinalities, which further exacerbates inter-method incompatibility and inconsistency [30, 42].

3.4.1.3 Crew activity primitive (CAP)

CAPs represent the base-level of task discretization in the HMT-centered HRA paradigm. These are fundamental actions taken in the course of achieving a major crew function (MCF)

[30, 87].

3.4.1.4 *Crew failure mode (CFM)*

CFMs represent the failure pathways that will result in a failed MCF, and therefore may lead to an HFE [30, 82]. These are related to the concept of proximate causes from the macrocognition framework, and are the crew-centered translation of individual psychological errors [2, 64]. Because the actualization of any CFM results in an MCF failure, CFMs can be seen as an analog to minimal cut sets in traditional reliability engineering block diagrams, and are mutually exclusive. For instance, the MCF "Identify failure (respond to alarms)" may have the following associated CFMs [84]:

- Miss alarm
- Incorrectly attribute alarm
- Dismiss alarm

3.4.1.5 *Major crew function (MCF)*

MCFs are the high-level actions of the HMT being studied, related to the Information, Decision and Action in Crew Context (IDAC) crew activities and macrocognitive functions [13, 30, 87]. MCFs are further decomposed into series of CAPs, and the different pathways by which HMTs can fail to accomplish a MCF are the Crew Failure Modes (CFMs) (see Table 3.1). The failure of any MCF could instantiate an HFE. As an example, the high-level objective *Respond to RCS Loop Thermocouple Failure* could be composed of the MCFs below [84]:

- Identify failure (respond to alarms) [I]
- Diagnose thermocouple failure [D]
- Take immediate actions [A]
- Enter correct procedure [A]

Note that MCFs here are built from macrocognitive functions, but represent a specific instantiation of a macrocognitive function applied to a system-related goal. MCFs are the composition of multiple CAPs.

3.4.1.6 Human failure event (HFE)

Human-influenced failures have traditionally been denoted in HRA models as human failure events, or HFEs. Because the understanding is now that errors can be initiated by either component of an HMT (i.e., the human or the system), HFEs should represent the failure of an HMT to complete a high level objective [30]. HFEs are decomposed into component MCFs, the failure of any of which might produce an HFE [30]. An HFE is therefore the end-result of the failure process, the failed state of some overarching objective which can involve multiple MCFs [88]. Human Error Probability (HEP), the probability of experiencing an HFE, should not be used to refer to the failure process itself.

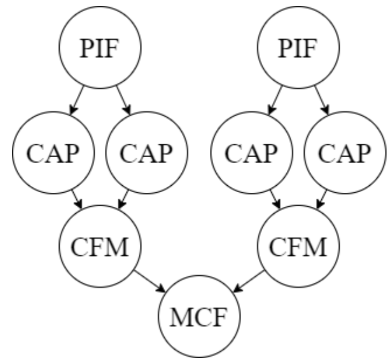
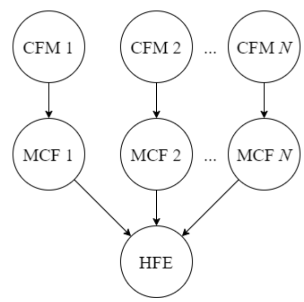
3.4.1.7 Human error probability (HEP)

The HEP is the probability of experiencing an HFE. Obtaining this value for use in a larger Probabilistic Risk Assessment (PRA) is generally regarded as one of the ultimate goals of any HRA. The HEP is not a variable or entity in its own right, it is the mathematical representation of the HFE.

3.4.2 Task Decomposition Terminology

The concepts defined below are meant to provide a general language for discussing and describing task decomposition (or task analysis) for HRA. These definitions provide a common understanding of the levels of a task decomposition across HRA. Previous research offers similar definitions for some of these concepts, particularly functions in the context of hardware [89–91]. Previous research on operator functional modeling has also explored the concept of a function for the HRA context [92–94]. The definitions provided here align generally with the concepts previ-

Table 3.1: Summary of the core HRA variables, their definitions and their relationships.

Concept	Definition	Visualization
Performance Influencing Factor (PIF)	The situational characteristics describing the context in which the Human-machine Team (HMT) performance occurs.	
Crew Activity Primitive (CAP)	The base-level task discretization representing fundamental actions taken to achieve a Major Crew Function (MCF).	
Crew Failure Mode (CFM)	The failure pathways that result in failure to achieve an MCF.	
Major Crew Function (MCF)	The high-level actions taken by a Human-machine Team (HMT); system-specific instantiation of macrocognitive functions.	(Adapted from [30])
Human Failure Event (HFE)	The highest-level failure considered in an HRA model. This is the culmination of the failure process that must include one or more *failed* Major Crew Functions, each of which is recorded as a corresponding Crew Failure Mode.	
Human Error Probability (HEP)	The probability of experiencing an HFE during a given scenario.	$HEP \equiv Pr(HFE \mid \{MCFs\})$

ously explored for functional modeling. However, the definitions offered in this section are meant to provide a general language for discussing the hierarchy of task decomposition constructs relative to each other, for instance the relationship (and distinction) between objectives, functions, and tasks.

3.4.2.1 [System] Event

A system event is not the goal of a crew action, but a milestone or system change that, when it occurs, represents a significant moment in the scenario response or operation. Events may be external to the HRA scenario, such as natural hazards, system failures or initiating events, or they may be catalyzed by the success or failure of an objective, i.e. an HFE. For instance, a system event can initiate or terminate an incident response, alter an HMT's operational posture, or signal a change in objective. In HRA, where various temporal discretizations may be applicable, system events may serve as useful discretization points to delineate objectives and response phases [95].

3.4.2.2 Objective

An objective in HRA should be considered the top level of task decomposition, which represents the ultimate goal of the portion of the event response under study. An objective could be the outcome of a procedure, or a set of procedure steps, or even a single step. An objective should be the top-most variable when discussing any distinct subset of HMT actions.

3.4.2.3 Function

A function is a high-level system purpose that the crew must achieve in order to successfully meet an objective, and is therefore effectively an abstraction of what the previously-defined MCF, which is a high-level system purpose that the crew must achieve in order to successfully meet an objective. Practically, the definition of a function will depend on how the objective is defined, but broadly the functions should be one "step" below objectives when discussing crew actions. For instance, if the objective describes the outcome of a set of procedure steps, each procedure step could be a function.

3.4.2.4 Task

A task should refer to a fundamental action on the system taken by an HMT which can be formulated from a verb-noun pairing, similar to a CAP [30, 83]. Another hallmark of a task is a concrete change in the scenario as a result of successful task performance. An information-gathering task results in additional information in the system, a diagnosis task produces additional knowledge, and an action task changes something in the system topology. Tasks should be defined as one "step" below functions when discussing HMT executions (i.e., decisions, actions, or communications). If the functions are defined at the level of procedure steps, tasks are portions in that procedure step (or procedure sub-steps). In HRA models, the task should be considered the lowest level of abstraction, which aggregate to form the higher-level concepts.

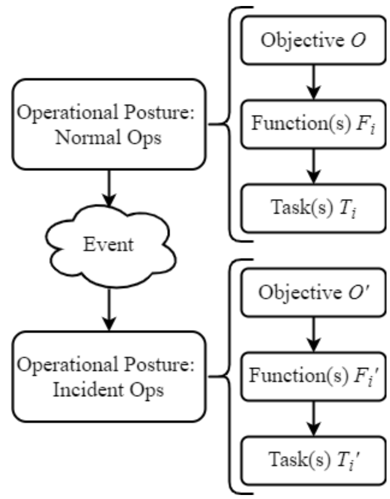
While tasks can, and should, be further decomposed for the purposes of human factors and ergonomics studies, HRA examines the integration and intersection of human and system activities, and is fundamentally a system-level analysis. Thus, HRA requires a higher level of granularity to produce interpretable results. HRA and human factors/ergonomics interrelate similarly to Newtonian Mechanics and Quantum Mechanics. Human factors and ergonomics is primarily interested in the "quantum" realm of subtasks and microtasks (i.e., how humans perform actions), whereas HRA looks at a more macro-level of HMT activity (i.e., how humans interact with eachother and the system). They are complementary fields, each useful in their own domain, but can cumbersome when used inappropriately.

3.5 Result: Mathematical Dictionary of Foundational Concepts

In addition to standardizing the definitions for fundamental constructs related to HRA, it is also imperative to unify definitions for key mathematical and statistical concepts as well. BNs, which are capable of explicitly modeling dependency relationships, form the basis of this dissertation in HRA, and are being widely investigated for their applications in HRA and PSA methods [30, 63, 64, 67, 77, 82, 86]. BNs are built on fundamental concepts in statistics, graph theory and set theory, so understanding HRA dependency and BN operation requires a thorough knowledge of these concepts [43].

Table 3.2: Summary of the event decomposition lexicography and their relationships.

Concept	Definition	Visualization
[System] Event	A milestone or system change which represents a discretization point; A significant moment in the scenario.	
Objective	Top level of abstraction referring to the ultimate goal of crew actions with respect to the plant/scenario need.	
Function	High-level purpose which crew must perform to achieve an objective.	
Task	Fundamental activity undertaken on system by the Human-machine Team; Lowest level of abstraction for HRA models.	



3.5.1 Set Theoretic Concepts

3.5.1.1 Union

The union of two sets X and Y is the set of elements belonging to either set. The union of two events is the situation where either or both events can occur [40]. The union of two events describes only the occurrence of either variable and does not imply any statistical dependence between the two. For example, understanding if an operator fails *either* the initial diagnosis (X) *or* the following immediate actions (Y) could be found as the union of the two events. In Boolean logic, the union is represented by "OR," "∪" or "+" [28]:

$$X \cup Y = \{i \mid (i \in X) \vee (i \in Y)\} = X + Y \tag{3.1}$$

3.5.1.2 *Intersection*

The intersection of two sets X and Y is the set of elements belonging to both sets. The intersection of two events is the incidence wherein both events occur [40]. The intersection of two variables describes only their simultaneous occurrence and does not imply statistical dependence. For example, understanding if an operator fails *both* the initial diagnosis (X) *and* the following immediate actions (Y) could be found as the intersection of the two events. In Boolean logic, the intersection is represented by "AND," "$\cap$" or "$\cdot$" [28]:

$$X \cap Y = \{i \mid (i \in X) \wedge (i \in Y)\} = X \cdot Y \tag{3.2}$$

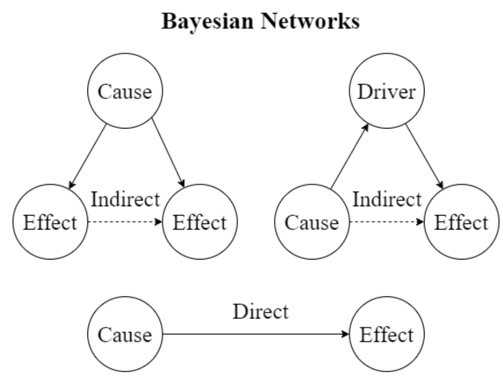
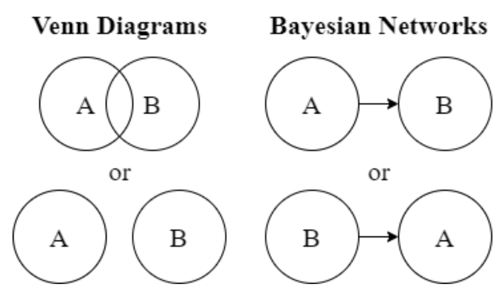
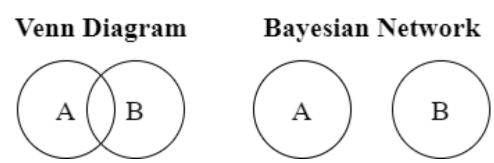
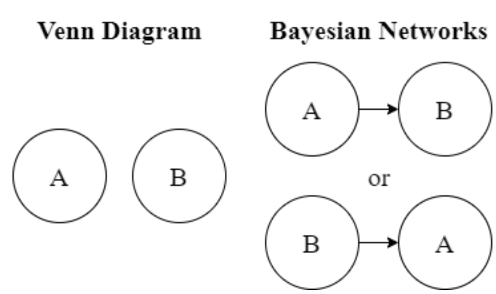
3.5.1.3 *Mutual Exclusivity*

Two sets X and Y are mutually exclusive (disjoint) if they have no common elements. Similarly, two events are mutually exclusive when they cannot occur at the same time [40]. For example, a diagnosis task cannot fail via both "misdiagnosis" and "no diagnosis" – the CFMs are defined as mutually exclusive. As discussed in Section 3.3.2, this is intimately related to dependence, and mutually exclusive variables are *not* independent. For mutually exclusive sets/events, the intersection is the empty set, $\emptyset$ [28]:

$$X \cap Y = \emptyset \Longleftrightarrow \{\nexists\, i \mid (i \in X) \wedge (i \in Y)\} \tag{3.3}$$

The above three concepts are taken from set theory, and are easily described using a Venn diagram (see Figure 3.4). It can be tempting to extend this visualization to express the independence of two events, however Venn diagrams are incapable of distinguishing independent and dependent events, as is shown in Table 3.3.

Table 3.3: Comparison of HRA dependency and statistical dependence concepts. Bayesian Networks are capable of explicitly distinguishing independent and dependent events, while Venn Diagrams are not.

Concept	Definition	Visualization
HRA Dependency	A dependency relationship exists between two HRA variables if they are connected by a direct or indirect causal relationship that changes the conditional probabilities of the variables; dependency exists regardless of whether the existence or utility of the variables is acknowledged within HRA.	
Statistical Dependence	Two events *A* and *B* are statistically dependent if the occurrence of one changes the probability of occurrence for the other. Mathematically, this means: $Pr(A \cap B) \neq Pr(A) \cdot Pr(B)$ $\implies Pr(A\|B) \neq Pr(A)$ $\implies Pr(B\|A) \neq Pr(B)$	
Statistical Independence	Two events *A* and *B* are statistically independent if the occurrence of one has no effect on the probability of occurrence for the other. Mathematically, this means: $Pr(A \cap B) = Pr(A) \cdot Pr(B)$ $\implies Pr(A\|B) = Pr(A)$ $\implies Pr(B\|A) = Pr(B)$	
Mutual Exclusivity	Two events *A* and *B* are mutually exclusive if they cannot exist simultaneously. Mathematically, this means: $Pr(A \cap B) = 0$ $\implies Pr(A\|B) = 0 \neq Pr(A)$ $\implies Pr(B\|A) = 0 \neq Pr(B)$ Note: this is also called "negative dependence," indicating that this is actually a type of dependence and *not* equivalent to independence.	

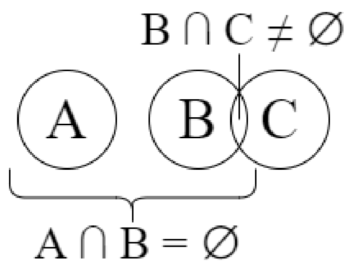


Figure 3.4: Venn diagram of three events. Events A and B are mutually exclusive, while events B and C have a non-empty intersection. Note that B and C may be mutually independent or dependent, there is no way to represent this relationship in a Venn diagram.

3.5.2 Statistical Concepts

3.5.2.1 Marginal Probability

An event A has marginal probability $Pr(A)$ or $f(A)$, which describes the probability of A occurring without considering the effects of any other event in the universe containing A. In practice, the effects of other events must be marginalized out, by summing or integrating the joint probability distribution function (PDF) with respect to the other events, to obtain a marginal probability [28, 43]. For example, given two events x and y with a joint PDF $f(x,y)$, the marginal PDF of x, $h(x)$, is obtained by integrating the joint PDF with respect to y [28]. For discrete variables, as are common in HRA, the marginal distribution $h(x)$ of a variable x would be obtained by summing the joint distribution $f(x,y)$ for all values of y. For example, finding the marginal HEP for a task subject to a single PIF with states $\{a,b,c\}$, can be found by summing the joint distribution over each of the PIF states.

$$h(x) = \int_{-\infty}^{\infty} f(x,y)dy \quad or \quad = \sum_{y_i} f(x, y = y_i) \tag{3.4}$$

3.5.2.2 Joint Probability

The joint probability of two events A and B is the probability that both A and B occur. For example, determining the probability that both a misdiagnosis (A) and a failed action (B) occur is equivalent to identifying the joint probability of both events. Mathematically, the joint probability, $Pr(A \cap B)$ or $Pr(A,B)$, is related to the marginal probability and conditional probability ($Pr(B|A)$) by equation 3.5 [43].

$$Pr(A,B) = Pr(A) \cdot Pr(B|A) \tag{3.5}$$

3.5.2.3 Conditional Probability

The conditional probability of an event is the probability of its occurrence, conditioned on the occurrence of another event in the universe. If two events are independent, then the conditional probability of one event, conditioned on the occurrence of the other, is equal to its marginal probability (i.e., does not change based on the other event). For a dependent event, the conditional probability is not equal to its marginal probability. Using the same example as above, the conditional probability of a failed action (B) given a misdiagnosis (A) is found as $Pr(B|A)$. The conditional probability illustrates the effect of the dependency between two events via the change in probability from the marginal value; it can be expected, for example, that a misdiagnosis increases the probability of failing immediate actions, and therefore that $Pr(B|A) \neq Pr(B)$.

3.5.2.4 Independence

Two or more events, in the same universe, are independent if the occurrence of one event has no bearing on the probability of occurrence of another event [43]. Mathematically, this means that the conditional probabilities of independent events are equal to their respective marginal probabilities [28]:

$$Pr(A|B) = Pr(A) \iff A \perp\!\!\!\perp B \tag{3.6}$$

Returning to Figure 3.4 and the representation of independent events, it is clear from equations 3.5 and 3.6 that the independence of two events with a non-empty intersection (e.g., events B and C in Figure 3.4) can not be determined. Venn diagrams can portray the independence or dependence between two events only in the case of mutually exclusive events (e.g., events A and B), which can be shown as completely dependent, because the occurrence of one event absolutely precludes the occurrence of the other. Then the conditional probability becomes zero, which is clearly not equal to the marginal probability except in the trivial case. Dependent events therefore do not follow Equation 3.6, but instead are defined by conditional probabilities which are not equal to their marginal probabilities:

$$Pr(A|B) \neq Pr(A) \iff A \not\perp\!\!\!\perp B \tag{3.7}$$

3.5.2.5 Bayes' Theorem

Bayes' Theorem uses observations (evidence, E) of an event or parameter θ to refine the assigned probability of that event. Prior to obtaining evidence, the probability of θ, $Pr(\theta)$, is assigned as the *prior probability*. Evidence is used to obtain the refined *posterior probability* via Bayes' Theorem, given by Equation 3.8 for the discrete case and Equation 3.9 for the continuous case [28]. Bayes' Theorem facilitates, among other uses, the Bayesian updating of baseline HEP values and PIF probabilities, and the relationship between the two, with new data for improving HRA, as shown in [63]. Practically, the conditional probability of the evidence E ($Pr(E|\theta)$) is found as the likelihood function of observing the evidence assuming the prior probability of θ. For instance, the conditional probability of making a diagnosis error given a PIF state would be found as the likelihood function of observing the misdiagnosis and the PIF state, normalized by the probability of the PIF state. Bayes' Theorem could then be used to update the HEP for misdiagnosis in light of the observed misdiagnosis (evidence).

$$Pr(\theta|E) = \frac{Pr(E|\theta) \cdot Pr(\theta)}{\sum_{i=1}^{n} Pr(\theta) \cdot Pr(E|\theta_i)} \tag{3.8}$$

$$\pi_1(\theta|E) = \frac{l(E|\theta) \cdot \pi_0(\theta)}{\int_{-\infty}^{\infty} l(E|\theta)\pi_0(\theta)d\theta} \tag{3.9}$$

3.5.2.6 Bayesian Networks (BNs)

BNs are a class of probabilistic graphical models that are capable of representing causal logic and dependency relationships [43]. A BN consists of a directed, acyclic graph with nodes and directed edges, and a conditional probability table (CPT) for each node [73]. BNs have several advantages relevant to HRA modeling: explicit representation of dependency relationships (in the directed edges); support of causal, evidential and intercausal reasoning between nodes; and multiscale, multimodal data fusion while maintaining traceability and computational simplicity [43, 63, 77, 96]. Table 3.3 illustrates the utility of Bayesian Networks over Venn diagrams for visualizing dependence and independence.

3.5.2.7 Evidence

Evidence for a Bayesian Network, in the HRA context, can come in different forms. Colloquially, "evidence" often refers to observations of specific states of HRA variables, such as the success or failure of an MCF, or the state of a PIF [63]. When this evidence is propagated through the network, the states of non-observed variables are changed via the encoded dependencies. "Hard" evidence describes the instantiated nodes, whereas "soft" evidence refers to the non-determined nodes with states updated by propagating evidence through the network [73]. Hard evidence can be further distinguished between "specific" and "virtual" evidence. Specific evidence refers to direct observations of the nodes, while virtual evidence elicits node state judgments from experts based on "undisclosed observations" [74]. In the most general sense, evidence is any information used to update the BN after initial construction.

3.6 Result: Robust Definition of Dependency in HRA

Chapter 2 illustrates that no comprehensive, robust definition of dependency exists in HRA methods. In addition, the myriad industry standards provide different, partially-correct definitions of dependency. Underlying this gap is the paucity of discussion over what dependency actually means in an HRA context [14]. The definitions that have been put forward are vague and limited to relationships between only certain variable types. However, dependency should be considered between and among all HRA variables. It is well-recognized that dependent tasks produce a much higher HEP; it should be clear that if other HRA variables can be dependent, ignoring the dependency will produce incorrect results. Therefore, it should be recognized that dependency must be considered between all HRA variables, and that HRA dependency is conceptually distinct from statistical dependence. Although the quantitative impact of HRA dependency is inherently linked to statistical dependence, it will be shown that statistical dependence is incapable of completely describing HRA dependency.

There is no *a priori* reason to restrict the conceptualization and application of dependency to tasks/HFEs in HRA. Limiting dependency in this way prioritizes pragmatic concerns of computational simplicity and expediency over the normative goals of accuracy and completeness. Deconstructing HMT performance into a web of discrete elements is also a simplification, but one which is necessary to develop diagnostic or prognostic capabilities. Within this simplified framework, every attempt should be made to retain remaining complexities and ensure the completeness of our models. The applicability of dependency between all HRA variables is a consequence of the reality in which HMT performance occurs. HMTs perform complex actions which take place in, and are (partly) determined by, a complex and dynamic situation, which includes personal, technological, team-based, organizational and environmental elements. Neither the actions of an HMT nor the factors influencing the actions exist within a vacuum. The relationships that link the elements in a performance model are representative of the cognitive, functional, physiological and/or phenomenological connections inherent in the situation under study. These relationships are present in every situation modeled in HRA; even in the trivial example of an HMT working towards a single objective affected by a single comprehensive influencing factor, the recognition that the objective is accomplished within some environment is necessarily indicative of dependency.

In the complex situation experienced by HRA, it should be obvious that any HRA variable (e.g., PIFs, CFMs, MCFs) can experience dependency [30]. If the HRA variables are viewed as

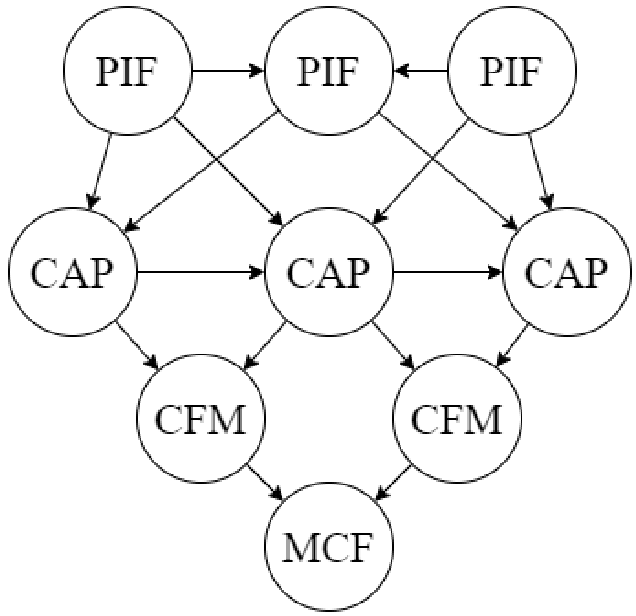


Figure 3.5: MCF Decomposition with dependencies (adapted from [30]).

elements in a hierarchy, e.g. the visualization in Table 3.1, dependency relationships can exist between variables in the same level or across adjacent hierarchical levels. Figure 3.5 illustrates possible dependency structures between HRA variables. For example, PIFs can experience dependency relationships with other PIFs or with variables in the next hierarchical level, which may be CAPs or CFMs depending on model construction. PIFs are the orthogonal discretization of a highly interdependent and dynamic situation defined at the intersection of the organization, technology and the environment. Accordingly, dependency should be viewed as an inherent property of the PIFs, derived from the physical and phenomenological relationships observed in reality. Appendix G of SPAR-H identifies possible PIF inter-relationships, but no guidance is provided on incorporating the dependencies into quantitative analyses, resulting in functional independence among the PIFs in SPAR-H [31].

Like PIFs, CFMs can also be subject to dependency. The set of CFMs corresponding to a single MCF represent orthogonal failure modes. As with PIFs, the orthogonality of the set should not be misconstrued as independence. Within a single MCF, the CFMs are mutually exclusive, and necessarily dependent; like a single component, a single function can only have one recorded failure mode. Dependency may also exist between CFMs across MCFs, meaning the observed

CFM of the antecedent MCF may change the probability of CFMs in the subsequent MCF. It should be noted that CFMs in separate MCFs are not a priori mutually exclusive. MCFs and HFEs are also subject to dependency, as is widely accepted in HRA. The recognition that dependency is an attribute of all HRA variables is critical to correctly constructing and evaluating HRA causal models.

3.6.1 HRA Dependency as Contrasted with Statistical Dependence

The explicit basis in phenomenological and/or physiological relationships differentiates HRA dependency from statistical dependence. Statistical dependence is a relationship between variables where the state of one variable alters the state probability of another variable [28]. Mathematically, this definition is complete for all contexts in that *any* two variables related in such a way are statistically dependent. These relationships are symmetric and devoid of any causal meaning, spawning the adage "correlation does not imply causation" [97]. Statistical dependence is a purely quantitative relationship between two variables, altering the conditional probabilities without expressing any causal connection between the variables. If two variables A and B are statistically dependent, it may be that A causes (directly or indirectly) B, or vice-versa, or they may share some common cause. For a normative, causal modeling-based field (i.e., HRA), these are genuine relationships which are required to understand the causal logic of the situation at hand. However, A and B may be causally-unrelated in reality, trending together by chance alone. Including such a spurious relationship in a causal model would be counter-productive and produce inaccurate results.

3.6.1.1 Sources of Statistical Dependence

It is clear that statistical dependence relationships have two possible sources: causality and coincidence. Causality, both direct and indirect, is the basis of all meaningful relationships in HRA. Coincidental relationships should be identified and eliminated from any HRA model. Simply modeling variables as dependent without verifying a causal connection is anti-productive to the mission of HRA. In addition to supporting the validity of the modeled relationship, the causal aspect of HRA relationships provides important contextual information to the model. In a normative context like HRA, a relationship's implication to the elements involved and to the broader system provide information as critical as the quantitative, probabilistic changes. For example, understanding the dependency between the PIFs "HMI Quality" and "Operator Stress" provides more information to the system than simply accounting for the increased probability of Operator Stress being rated "High" when HMI Quality is "Poor." Understanding the implication of this dependency, that HMI Quality can affect Operator Stress but not vice-versa, provides additional knowledge about the system operation that can facilitate design, operation and analysis decisions affecting reliability. Using statistical dependence as a stand-in for dependency relationships in HRA would lose the "implicational" aspects of the relationships and produce inaccurate results from the model. Dependency relationships in HRA must be rooted in causality, to avoid including coincidental relationships. When implemented in a scenario model, coincidental relationships connect variables that have no physical, functional or situational association. Including them would inaccurately characterize the probabilities of the affected variables, create inaccurate models of the situation and invalidate the analysis.

3.6.2 Requirements for a Robust Definition of Dependency

To ensure a firm foundation for the future of dependency study in HRA, this dissertation creates a concise, comprehensive definition for what constitutes dependency in HRA. To correct the shortcomings of previous attempts at a defining dependency, a comprehensive definition must explicitly include:

1. A basis in causality to prevent the inclusion of spurious relationships,
2. A probabilistic impact that follows the form of the "quantitative imperative" of statistical dependence,
3. Variable-agnostic language so that the definition is applicable to all HRA variables.

3.6.2.1 *Causal Basis*

Causality, rather than coincidence, must be the progenitor of HRA dependency if the goal is to determine the effects of *real* relationships between HRA variables. Current HRA dependency assessments rely on checklists of coincidental factors (e.g., temporal proximity, physical proximity, crew similarity) to determine a level of dependency between two tasks or functions [16, 31]. However, basing dependency in causality may not be as far a logical leap from current HRA practice as it may seem. The logical basis for the selection of dependency factors must be an implicit sense of causality; as Section 3.6.1.1 shows, the only alternative basis for dependency assessment would be coincidence (in the colloquial sense).

Spurious relationships, i.e. those with no determinable basis in physical or phenomenological coupling, are anti-productive to system modeling. Without an underlying causal connection, using such relationships as the basis for HRA dependency serves to change variable state probabilities and, therefore, the outcomes of the HRA and parent PRA, without considering the causal context. This lowers the realism of the HRA assessment in the pursuit of conservatism. However, if the THERP-derived dependency levels (Section 2.4.1, Figure 2.2) are incorporated, this conservatism is lost due to the low-end biasing of the THERP method (Section 1.2.1). There is, therefore, an urgent need to move on from the coincidence-based paradigm established by THERP to a causally-oriented basis for assessing dependency relationships.

3.6.2.2 *Explicit Probabilistic Nature*

HRA, as implemented within a typical PRA process, is necessarily a probabilistic discipline. The overall goal of most HRAs is to produce the HEP, the human error *probability*. Dependency is similarly probabilistic in nature, as even current implementations center around updating HEPs in light of inter-task relationships. However, the probabilistic nature of dependency is not always communicated clearly in the definition. Definitions of dependency vary in their presentation of the probabilistic aspects of dependency, ranging from incorrect suggestions that dependency "may" change the probabilities involved to more appropriate recognition that the probability of at least one involved variable *must* change in the dependency relationship.

Because HRA dependency is rooted in statistical dependence, the recognition that probabilities are changing is a necessity for understanding the relationship, and therefore for assessing the relationships in any HRA method. Accordingly, a robust definition of dependency must explicate the probabilistic nature of dependency *within* the definition itself.

3.6.2.3 Variable Agnostic Application

As discussed in Section 1.2, definitions of dependency hitherto relied upon are largely restricted to addressing HFE-HFE dependency. This limitation presents a problem particularly when dealing with PIFs, which are necessarily interdependent, and generally neglects the complex reality underpinning why HRA is performed. Statistical dependence is a relationship that applies between any two variables ([28]), and the same is true for HRA dependency. There is no *a priori* reason to limit the definition of dependency, particularly when relationships between other variables are known to exist (see SPAR-H Appendix G, [31]). Dependency between HFEs is known to have significant effects on the overall results of the HRA (i.e., the HEP) - including other dependency relationships is likely to have a similar effect. More concerning, *neglecting* these other dependency relationships may be biasing HRAs by not accounting for relationships that increase the probabilities of negative variable states (e.g., HEP).

A robust definition of dependency should, therefore, not be limited to relationships between only one variable type, nor to a single type of relationship. As Chapter 4 demonstrates, there are multiple types of relationships possible between multiple variables.

3.6.3 Robust Definition of Dependency

With the requirements from Section 3.6.2 at the forefront of consideration, HRA dependency can be comprehensively defined as follows:

A dependency relationship exists between two HRA variables if they are connected by a direct or indirect causal relationship that changes the conditional probabilities of the variables; depen-

dency exists regardless of whether the existence or utility of the variables is acknowledged within HRA.

3.7 Discussion: Definitions for a Comprehensive HRA Foundation

The definitions provided in this chapter provide a standard and universal set of HRA concepts and modeling constructs that strengthen the theoretical and technical basis of HRA. The lexicographical definitions created in this chapter combine the amenable portions of the defined concepts from HRA literature, while also providing new information that serves to more robustly define the concepts for immediate use in HRA. The definitions for mathematical concepts incorporate the standard, accepted definitions, which can be found in introductory texts on mathematics and statistics and/or published sources of HRA, and refines the definition with a focus on the HRA context. This chapter therefore presents a novel comprehensive collection of definitions for standardizing the foundations of HRA. Standardization will reduce the prevalence of misinterpreted or incorrectly-used concepts in HRA (e.g., the use of HFE as interchangeable with task, function, or objective).

As implied in the definition requirements, the definition of dependency provided here does not change the quantitative meaning of dependency in HRA, which is identical to statistical dependence (see Equation 3.7). HRA dependency is statistical dependence reinforced with the causal implication of the real-world functional relationship between the variables. This definition of HRA dependency is compatible with the current work being done on statistical dependence in HRA [98, 99]. The addition of causality serves to add a basis and interpretability to HRA relationships. The definition of dependency does not indicate that every dependency relationship will produce the same quantitative effects, only that Equation 3.7 provides an equation form which all dependency relationships will follow; the magnitude of change will depend on the specific relationship.

3.8 Conclusions

This chapter defined many constructs and concepts that are critical to understanding, modeling, and quantifying HRA dependency in an objective, traceable, and accurate manner. The definitional basis thus provides a robust and uniform technical basis for the continued improvement of HRA as a field. The results of this chapter will help to reverse the effects of the previously-siloed approach to HRA development, by allowing methods and analysts to intercommunicate and build a more comprehensive understanding of why, how, and how often human error occurs. This chapter also provides the tools to build HRA models from a standard set of modeling constructs, meaning the next generation of HRA methods will be traceable and reliant on a consistent and understood set of concepts and constructs. As a result, this chapter addresses the lack of a uniform definitional basis for HRA, and thereby closes the first technical gap identified in HRA dependency.

The HRA constructs/variables defined in this chapter are the components necessary to building comprehensive causal HRA models and ultimately obtain a more comprehensive understanding of human reliability. Further, the robust definition of HRA dependency and the defined modeling constructs provided in this chapter provide the foundation for the remainder of this dissertation on conceptualizing, modeling, and quantifying HRA dependency. The robust definition of dependency is operationalized for HRA modeling with the development of the HRA dependency idioms in Chapter 4.

Chapter 4: Developing the Graphical Structure of HRA Dependency: HRA Dependency Idioms

This chapter addresses Research Objective 2, the development of the fundamental relationships and the corresponding graphical structures describing dependency in HRA. Framed as HRA dependency idioms, these structures draw on extant notions of causal structuring in Bayesian Networks and connect them to the HRA variables described in Chapter 3 to develop a set of distinct HRA relationships and their accompanying BN structures.

This chapter is reproduced from material under review in the journal Nuclear Science and Engineering.

4.1 Introduction

The development of an accurate, quantitative human reliability analysis (HRA) method requires a thorough understanding of how human-machine teams (HMTs) perform and fail actions. HMT actions and performance influencing factors (PIFs) do not exist in a vacuum, and are thus highly dependent on each other. The original HRA method, THERP, recognized this aspect of main control room (MCR) operations, and included guidelines for assessing a qualitative level of dependency between two tasks, as well as equations for modifying the human error probability (HEP) of a task on the basis of the dependency [16]. The most widely-implemented HRA methods in nuclear applications derive aspects from THERP, and still use the same architecture to account for dependency between tasks [31, 39]. Dependency analysis identifies and models the reasons that two HRA constructs (e.g., tasks or human failure events, HFEs) are interconnected. Quantifying dependency imposes a significant change to the final human error probability (HEP) [15, 44]. Most current dependency frameworks draw heavily on the THERP dependency framework, which centers around identifying correlation-based similarities between two tasks. Such methods therefore do not explicitly base dependency on causal relationships, and limit the scope to only dependency between tasks or HFEs. As a result, current dependency analysis techniques impose significant changes to the results of HRA (and therefore probabilistic risk assessments) without modeling the reality of a scenario or basing the probability changes on data or literature. This chapter will show that these aspects are inadequate to appropriately model the complex reality of HRA dependency, and creates a new framework to conceptualize and model dependency relationships.

The current understanding of dependency is largely limited to direct, correlational relationships *only between HFEs* (e.g., 'error begets error'), and is supported by the widespread implementation of THERP descendant methods [12, 17, 18]. Implementing an expanded framework that includes more complex relationships between all HRA constructs will produce more accurate models of HMT task performance and more appropriate HEP estimates. Chapter 3 developed a unified definitional basis to address the first of these issues and set the foundations of this chapter. As shown in Chapter 2, current research efforts allay immediate concerns regarding dependency usability and traceability, but do not address all of the technical gaps relevant to the foundations of HRA dependency. This chapter creates a set of HRA dependency idioms (archetypical relationships) that implement the probabilistic, causal definition of dependency created in Chapter 3. The present work fills long-standing technical gaps in the field by addressing more complex relationships between HRA variables beyond the HFEs, assessing dependency in a causal framework, and providing a formal modeling structure for HRA dependency.

4.2 Approach: Bayesian Networks for HRA Dependency

This dissertation asserts that dependency should be viewed more comprehensively, to include additional relationships that can exist between HRA elements, as described in [30]. Additionally, and perhaps more critically, the conceptualization and implementation of dependency in HRA must be grounded in causality through a basis in physical, phenomenological and/or cognitive relationships. Chapter 3 focused on providing definitional clarity to dependency and surrounding concepts, and ensuring that dependency in HRA is causally-based, probabilistic, and applicable beyond HFE-to-HFE relationships [44]. This section reviews why BNs are a suitable architecture for visualizing and modeling dependency relationships based on this definition.

4.2.1 HRA Constructs

Previous work using BNs for HRA developed a comprehensive set of HRA constructs to encode in the BN architecture, as well as the resulting dependencies that must be included [30]. The HRA constructs include the PIFs, Crew Failure Modes (CFMs), Major Crew Functions (MCFs) and the Human Failure Events (HFEs); these are described in detail in Chapter 3 [30, 44]. Briefly, the PIFs are the multi-dimensional factors that will cause changes in human-machine team performance ([42]), and take informative states relating to specific characteristics of the given factor. CFMs are the mutually exclusive ([44]) ways in which an MCF can be failed, and take informative states relating to the specific mode of MCF failure. MCFs are the tasks/functions being performed, and will take binary, *uninformative* states that relate only whether there was a success or failure, with no information on the mode of failure [44]. HFEs are defined by one or more MCF failure, and will take similarly *uninformative* states that relate only whether an HFE occurred or did not occur. Using these HRA constructs, Groth, et al. (2019) identified the following dependency relationships that must be included and quantified in an HRA model [30]:

$$Pr(PIF_a|pa(PIF_a)) \tag{4.1}$$

$$Pr(CFM_b|pa(CFM_b)) \tag{4.2}$$

$$Pr(MCF_c|pa(MCF_c)) \tag{4.3}$$

$$Pr(HFE|MCFs) \tag{4.4}$$

In the dependency relationships given above, $pa(X)$ refers to the "parents" of node X, which are all those nodes that have a directed arc leading into node X. There are then at least four classes of conditional probabilities (i.e., dependency relationships) that must be accounted for in any dependency model. Of course, other nodes may be included (e.g., Crew Activity Primitives (CAPs) that are task-level elements), which then increases the number of corresponding dependency relationships. This dissertation focuses on the HRA constructs most readily available in HRA data, which are chiefly the PIFs, MCFs, and HFEs defined in Chapter 3.

The elements and relationships given in [30, 44] represent critical constructs for HRA models that must be included as nodes in a Bayesian network for HRA. The causal relationships between the constructs are encoded on the BN as directed arcs. The directed arcs model the causal dependencies between HRA elements, and also determine the pathways for information propagation around the network [43].

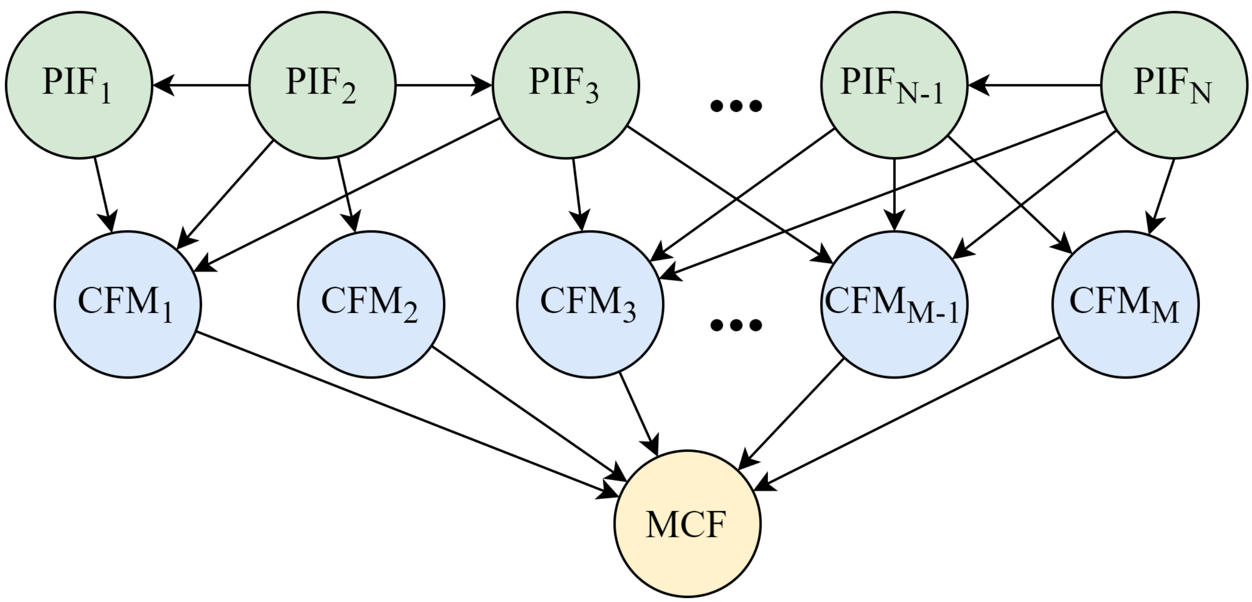


Figure 4.1: Semi-hierarchical BN created from HRA elements [30].

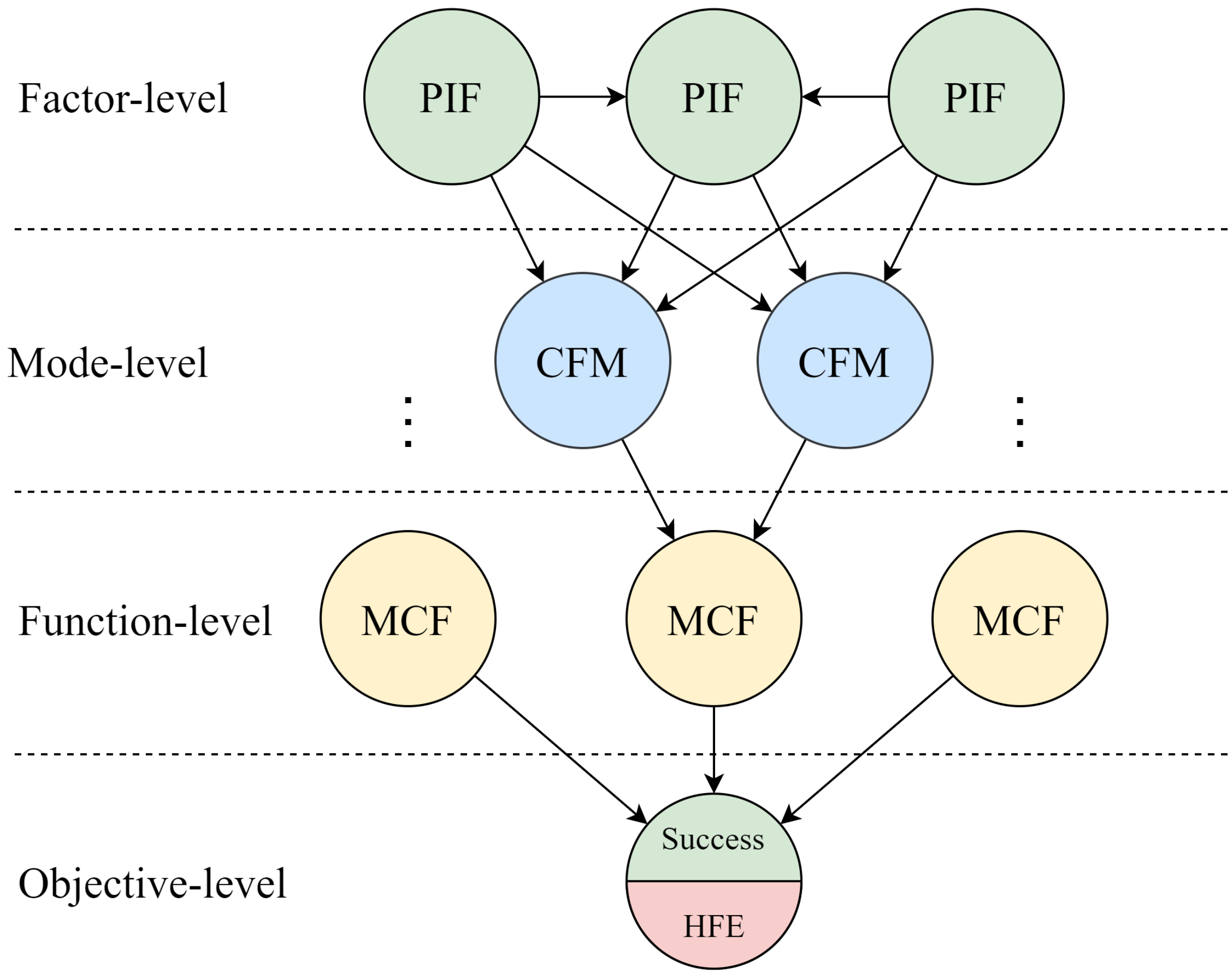


Figure 4.2: Expanded semi-hierarchical BN created from HRA elements. Adapted from [88].

4.2.2 Bayesian Networks for HRA

4.2.2.1 HRA Constructs on BNs

Encoding the principal HRA constructs in a BN creates a semi-hierarchical structure as shown in Figures 4.1 and 4.2. This structure arranges the HRA constructs ([30]) in a logical manner ([88]) and imposes conditions on the dependency relationships that are *possible* in the model. As can be seen from Figure 4.1, not all variable types are interconnected. Within the span of one MCF (i.e., within a single "timestep"), PIFs can be connected to PIFs and CFMs, CFMs are connected to MCFs; MCFs connect to HFE nodes. There are similar restrictions on the relationships possible across MCFs (i.e., connecting different tasks/functions across time). Across MCFs, PIFs can interact with other PIFs and CFMs, and CFMs from one task can affect future PIFs and CFMs, as shown in Figure 4.3.

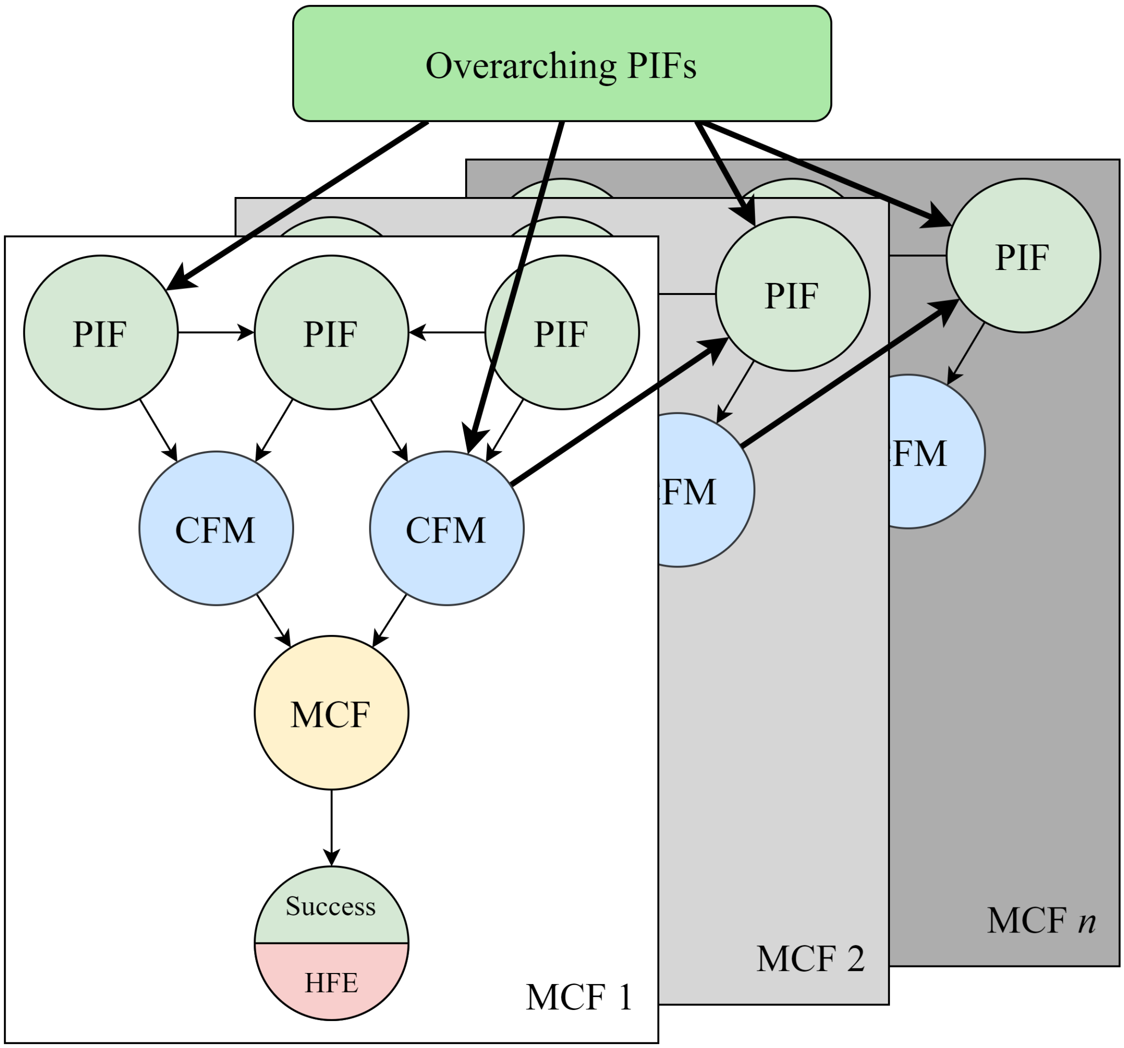


Figure 4.3: The semi-hierarchical BN architecture guides dependencies both within a single MCF and across MCFs.

The restrictions/conditions on dependency imposed by the BN structure are extended by the definition and implementation of the HRA constructs. Particularly, the notion of informative and uninformative states (i.e., CFMs and MCFs) adds a pragmatic restriction to modeling dependency relationships. The information contained in an MCF node, when CFMs are available/modeled, is relatively minimal – the MCF node encodes *only* whether the MCF was a success or failure; any information related to the mode/type of failure is contained in the associated CFM nodes. If the failure of an MCF will affect a future PIF (e.g., if "Misdiagnosis" reduces the "Available Time"), it is not simply the failure that enforces this dependency, but the *mode* of failure encoded by the CFM Misdiagnosis. The lack of dependency-relevant information encoded in the MCF nodes means that CFMs, not MCFs, will drive dependencies between tasks. Consider again the example of a crew performing "Diagnose Stuck Rod" (Task A) and "Enter Procedure SR-1" (Task B). In this example, identifying that Task A is failed conveys much less information than identifying the mode of failure (e.g., Misdiagnosis or Delayed Diagnosis). In addition, the mode of failure will have implications for dependency; a delayed diagnosis may reduce the time available and increase stressors to enforce a dependency on Task B. A misdiagnosis, on the other hand, will likely further reduce the time available (for recovery and Task B), and/or directly influence Task B failure modes (e.g., preventing performance).

4.2.2.2 Dependency on BNs

Dependency relationships between the HRA constructs are encoded as directed arcs in the BN. The directed arcs indicate the direction of causality, although information and reasoning can travel in both directions on a directed edge. This enables causal, evidential, and inter-causal reasoning between the HRA constructs [43, 97]. This is the basis for the "*d*-Separation" concept in which the reasoning path between two nodes is severed by conditioning (adding information) on a third node [74, 97].

The flow of information in a causal BN depends on the knowledge supplied to the graph and the structure of the graph itself (i.e., orientation of directed arcs). In Figure 4.4, knowledge about the state of node D or E opens the reasoning pathway between nodes A and B. With no knowledge about D or E, nodes A and B are independent. Thus, nodes sharing a common effect are only dependent if the common effect is known. Nodes C and D share a common cause in B and are dependent *without* knowledge of B. Conditioning on node B forces conditional independence between C and D [43]. Knowledge changes the pathways on which information, and dependency, can travel, but knowledge does not induce or break causal relationships; conditioning on nodes in the graph does not create new causal relationships. For HRA, dependency relationships are those relationships that *a priori* transmit causality, that is when *no information/evidence* is imparted to the graph. Thus, common-effect relationships are not dependencies in HRA (i.e., A and B in Figure 4.4 are not dependent), while common-cause relationships *are* dependencies. Claiming a dependency between two events that share a common cause is a regular practice, and using artificial common-cause nodes has been used to account for generalized dependency in graphical models [96].

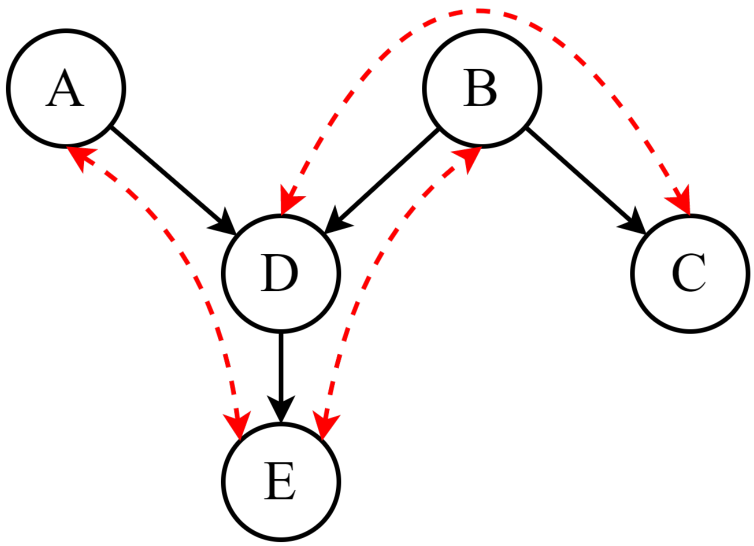


Figure 4.4: Causal BN framework. Directed arcs are shown with solid black arrows. Active reasoning pathways under no conditioning are shown with dashed red arrows. Note that A and B are not connected via an active path; Common-effect relationships are not dependencies in HRA, as they require evidence on a third node (e.g., D) to transmit the relationship.

The semi-hierarchical structure of the HRA BN, and the information content of the different nodes, are therefore critical to determining how dependencies are represented on BNs. This is true both when investigating dependencies that occur contemporaneously and those that exist across MCFs (i.e., across time). Figure 4.3 shows how the BN semi-hierarchical structure facilitates certain dependencies, both within a single MCF and across MCFs.

4.2.3 Existing Bayesian Network Idioms

Bayesian networks are used to build the dependency framework because of their ability to explicitly model dependency relationships between nodes. The rigorous BN structure, HRA constructs and modeling conditions (Section 4.2.2) mean that there are a finite number of edges in an HRA BN, and thus repeated causal structures are expected. Fenton and Neil identify four of these repeated structures and define them as *idioms*. The idioms are therefore repeatable graphical structures that describe specific relationship types. They identify four general idioms as fundamental causal structures, independent of application. The idioms defined in [43] are visualized in Figure 4.5 and explained below.

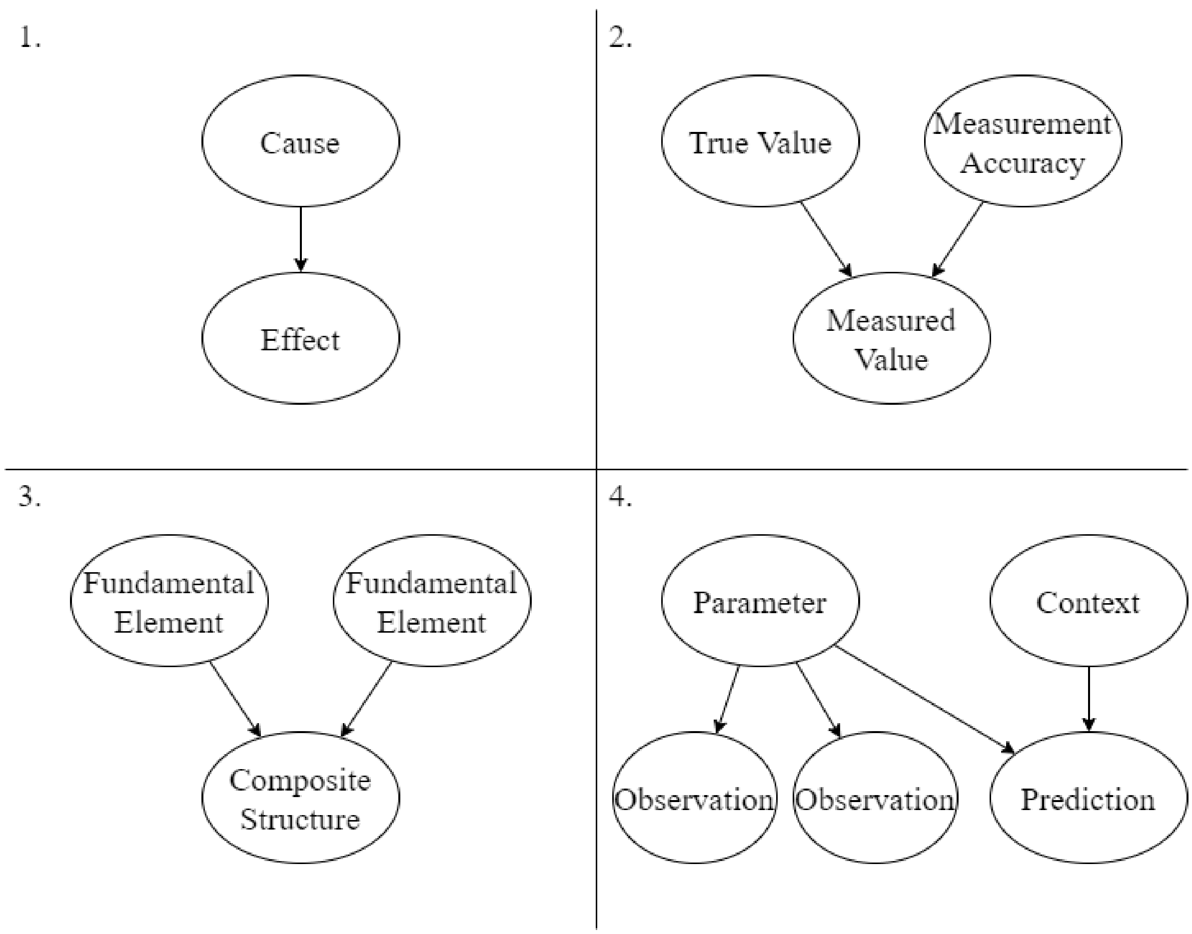


Figure 4.5: Bayesian network idioms defined in [43]. 1. Cause-Effect; 2. Measurement; 3. Definition/Synthesis; 4. Induction.

1. *Cause-Consequence:* Causal relationship between two nodes. The observation of parent node state (cause) changes the probability of observing a given child node state (effect).

2. *Measurement:* Relationship between a parameter's measured value, the measurement accuracy and the parameter's true value. The measured value is dependent on both the true value and the accuracy of the measurement.

3. *Definitional/Synthesis:* This is *not* a causal reasoning structure, but a hierarchical structure showing the decomposition of a composite node into its constituents.

4. *Induction:* Relationship between the observations of a parameter, the context of the observation, and a future prediction (under different context). Previous observations of a parameter can be combined with the new context to predict a new value.

In addition to the idioms proposed by [43], there are two conditional reasoning structures that should be mentioned with the idioms as fundamental graphical structures. The *conditional dependence* and *conditional independence* structures, shown in Figure 4.6, are reasoning relationships that can connect three nodes in a generic BN, and explain the concept of *d*-Separation [100]. Conditional dependence (Figure 4.6A) describes common-effect relationships wherein the parent causes are dependent *only* if evidence is available on the effect; this is not a dependency for HRA. Conditional independence (Figure 4.6B) describes common-cause relationships, wherein the effect nodes are dependent *unless* evidence is available on the cause.

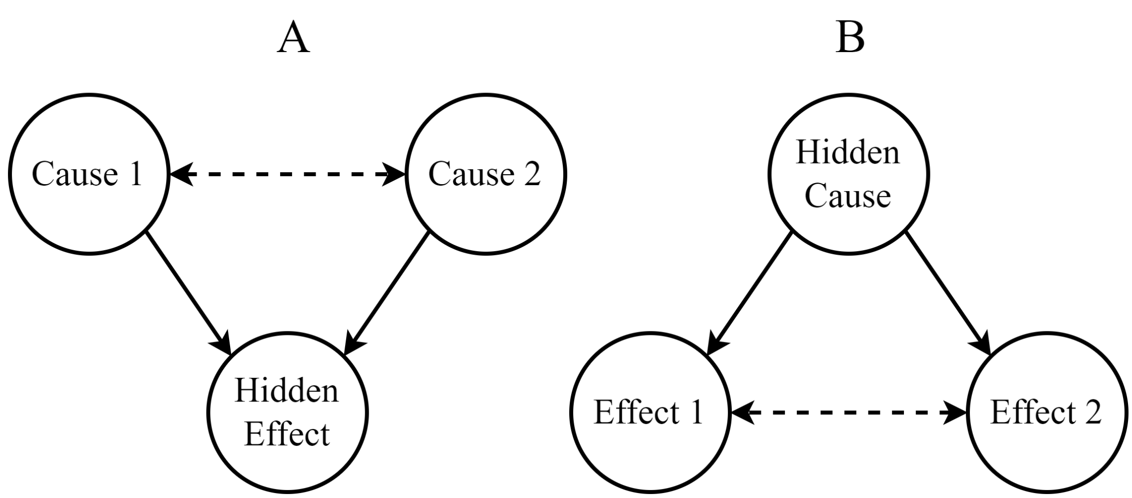


Figure 4.6: (A) Conditional dependence between Cause 1 and Cause 2; (B) Conditional independence between Effect 1 and Effect 2. Adapted from [100].

Both the qualitative and quantitative aspects of dependency relationships are critical to HRA. The qualitative information in the structures captures the causal information and the relationship type; the quantitative aspects indicate the impact (strength) of the relationship. Both are required to build realistic, accurate HRA models that are useful in diagnostic and prognostic applications. Dependency relationships are conferred by chain (e.g., A→D→E) or diverging (e.g., D←B→C) connections. These correspond to active reasoning pathways in the network when it is not conditioned on any nodes. This means that evidence added to the graph can break dependency relationships, but not create them. Simply put, dependency is defined on a causal BN as follows:

A dependency relationship exists between two nodes A and B if there exists an active path between A and B when conditioning the graph on an empty set of nodes.

The four idioms from [43] (Figure 4.5) and the conditional reasoning structures from [100] (Figure 4.6) illustrate generic causal logic structures for BNs. These structures informed the development of the HRA dependency idioms, which define relationships of interest to HRA.

4.3 Method: Creating HRA Dependency Idioms

This section details the creation of the HRA dependency idioms. The idioms represent archetypical dependencies for HRA, each of which are associated with distinct modeling structures built with the directed arcs in the causal BN architecture. Each dependency idiom is distinct, but applicable to multiple different HRA variables. These idioms were developed to describe and model the dependency relationships identified in [30], in accordance with the modeling limitations discussed in Section 4.2.2. This section will continue to consider the example of a crew performing Diagnose Stuck Rode (MCF A) and Enter Procedure SR-1 (MCF B) that will be referred to as the "Example Scenario." In the Example Scenario, it will further be assumed that both MCF A and MCF B are subject to the PIFs Workload, Time Available, Stressors, Training/Experience, and Fitness for Duty.

4.3.1 Causal

The causal idiom describes cause-effect relationships between HRA constructs, wherein one element causes another. This is the Cause-Consequence idiom from [43]. For HRA, and engineering in general, the most salient causation/causality paradigm is probabilistic causality, which embraces the inherent stochasticity in physical processes and causal mechanisms [101]. In probabilistic causality, causes are not invariably followed by their effects, but change the *probability* of the effect [74]. Although humans often fail to recognize the stochasticity of their environment, this should not be confused with the non-existence of stochasticity in the environment [14]. For example, smoking does not invariably lead to cancer, but it is still valid to state that "smoking causes cancer" because there is a probabilistic causal relationship such that smoking increases the probability of developing cancer. Similarly, high stress will not invariably result in a failed decision task, but it will increase the probability of failure, and therefore it is valid, in the frame of probabilistic causality, to say that "high stress causes failed decisions." Causal dependencies are transmitted through *causal mechanisms*, which are physical, phenomenological, or cognitive processes that exist to connect the parent and the child. For example, the mechanism in the causal dependency "high stress causes failed decisions" is likely the effect of increased stress hormone concentration on cognitive function [71].

Causal dependencies can occur between a number of HRA variables. PIFs can cause changes in other PIFs and in CFMs; CFMs, in turn, can cause changes in the PIFs surrounding *future* MCFs. As previously discussed, MCFs do not generally propagate causal dependencies because of their uninformative states, particularly when CFMs are present in the model. This is a pragmatic simplification, rather than a theoretical limitation. In the Example Scenario, causal dependencies will occur between the PIFs and CFMs for both MCF A and MCF B. For instance, High Workload may cause Delayed Diagnosis on MCF A, and High Stress may cause Reading Error on MCF B. Training/Experience will cause changes in the states of PIFs Time Available and Stressors.

Probabilistic causality, the framework incorporated in the present research, does not require

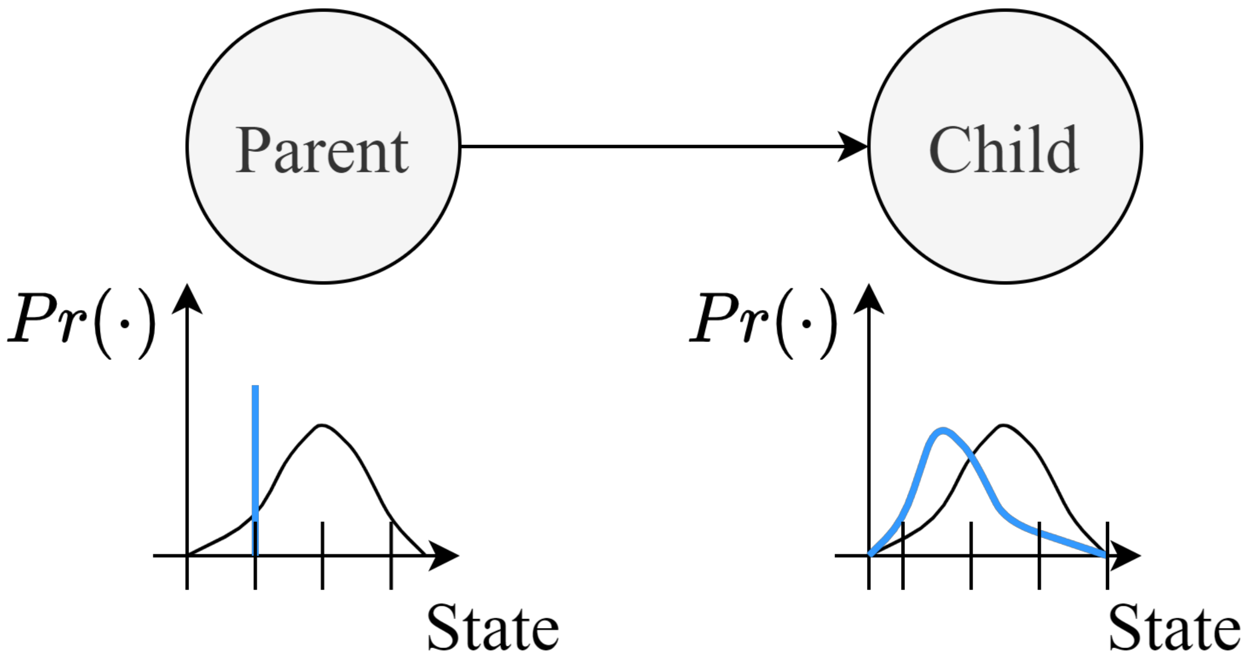


Figure 4.7: Causal dependency idiom between a parent node and child node. The black distribution represents a nominal distribution of node states. The blue line on the parent indicates an observation of a given state, which updates the child node state distribution to the blue curve.

the assumption of a causal mechanism as strictly as other causal frameworks, e.g., mechanistic causation [101]. However, model realism demands that a viable causal mechanism be proposed as the progenitor of dependency relationships in HRA, to avoid incorporating spurious dependencies and elucidate hidden confounding variables. Because causal mechanisms are construct-specific, it is beyond the scope of this dissertation to identify or propose causal mechanisms that generate each dependency idiom. The identification of causal mechanisms in HRA is the subject of both previous and ongoing research efforts [2, 64].

4.3.2 Definitional

There are dependencies in HRA that describe *definitional*, rather than causal, relationships. In definitional dependency, it is more accurate to say the parent defines or composes, rather than causes, the child. This is similar to the Definition/Synthesis idiom defined by [43] (Figure 4.5). There are significant implications for understanding and modeling the difference between definitional and causal logic. There is no causal mechanism in definitional dependency. Further, definitional dependency is deterministic, as opposed to the probabilistic nature of causal dependency, meaning the parent node state *precisely* determines the child node state, as Figure 4.8 shows. Moreover, the parent node states can be represented as distinct states of the child node when represented on a Bayesian Network.

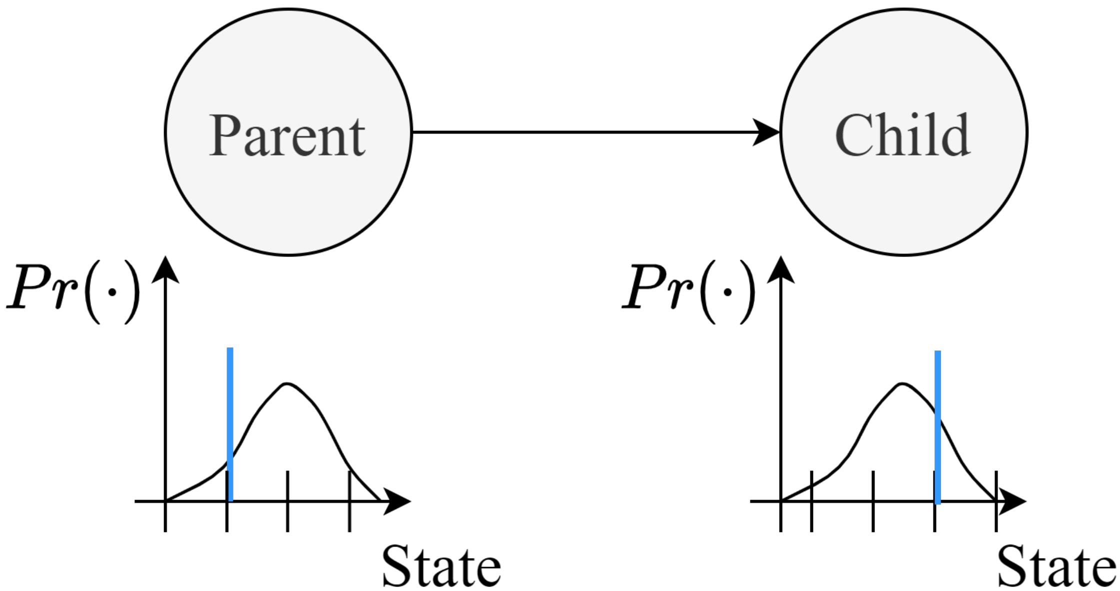


Figure 4.8: Definitional dependency idiom between a parent node and child node. The black distribution represents a nominal distribution of node states. The blue line on the parent indicates an observation of a given state, which precisely determines the child node state (blue line).

Definitional dependency is rooted in mathematical formalism – the relationship between sets and subsets – rather than causal logic. A CFM is a subset of the possible failure states of an MCF (which can be successful or failed via one of the MCFs). As a result, the observation of a CFM *necessarily* means that the MCF is failed. Thus, a CFM defines the failure of an MCF.

Definitional dependency is predominant between Major Crew Functions (MCFs) and their associated Crew Failure Modes (CFMs). Saying a CFM causes its MCF to be failed is not strictly accurate, as there is no discernible causal mechanism behind the relationship. Rather, a failure mode defines how the failure mansifests, that is the failure typology. For example, the hardware failure mode "premature operation" does not *cause* the failure (of, e.g., a pump), it defines *how* the failure occurred. Equivalently, the failure mode defines what type of failure occurred. In the

Example Scenario, the CFM misdiagnosis does not cause the crew to fail the MCF Diagnose Stuck Rod. A misdiagnosis defines *how* the crew failed the diagnosis MCF. The definitional dependency between CFMs and MCFs means that CFMs can be represented as distinct states of the MCF. An MCF with states {"Success" and "Failure"} informed by definitional dependencies from distinct CFM nodes could be modeled by a single MCF node with states {Success, CFM 1, CFM 2, etc.}. Therefore in the Example Scenario, MCF A could be represented with states {Success, Misdiagnosis, Delayed Diagnosis}. Due to the definitional dependency between CFMs and MCFs, MCFs can replace CFMs in any other dependency relationship. For example, a PIF-CFM (e.g., Time Available → Misdiagnosis) causal dependency could be replaced by a PIF-MCF causal dependency (e.g., Time Available → MCF A) if the states of the MCF correspond to the different CFMs.

Definitional logic may underpin relationships between other variables, such as between PIFs that are decomposed into constituent parts that define the aggregate node. However, definitional logic is predicated on certainty of the relationship, which is a rare feature of complex engineering systems. The Construction idiom describes seemingly-definitional relationships that are uncertain, such as inter-PIF relationships.

4.3.3 Construction

Construction dependencies describe uncertain definitional relationships. In construction relationships, the delineation between causation and definition logics may be blurred, and seemingly either conceptualization suits the relationship. It may be equally valid to state that the parent node causes the child, and the parent node defines the child. However, due to the (epistemic or aleatory) uncertainty in the relationship, the parent node state does not precisely determine the child node state, as shown in Figure 4.9.

Construction dependency is distinguished from Causal dependency by the aspect of definition in the relationship: the parent nodes may be said to define the child. Construction is also distinct from Definitional dependency due to the incorporation of both the element of uncertainty and the probabilistic aspect of causality. Nodes defined by construction dependency may be consolidated

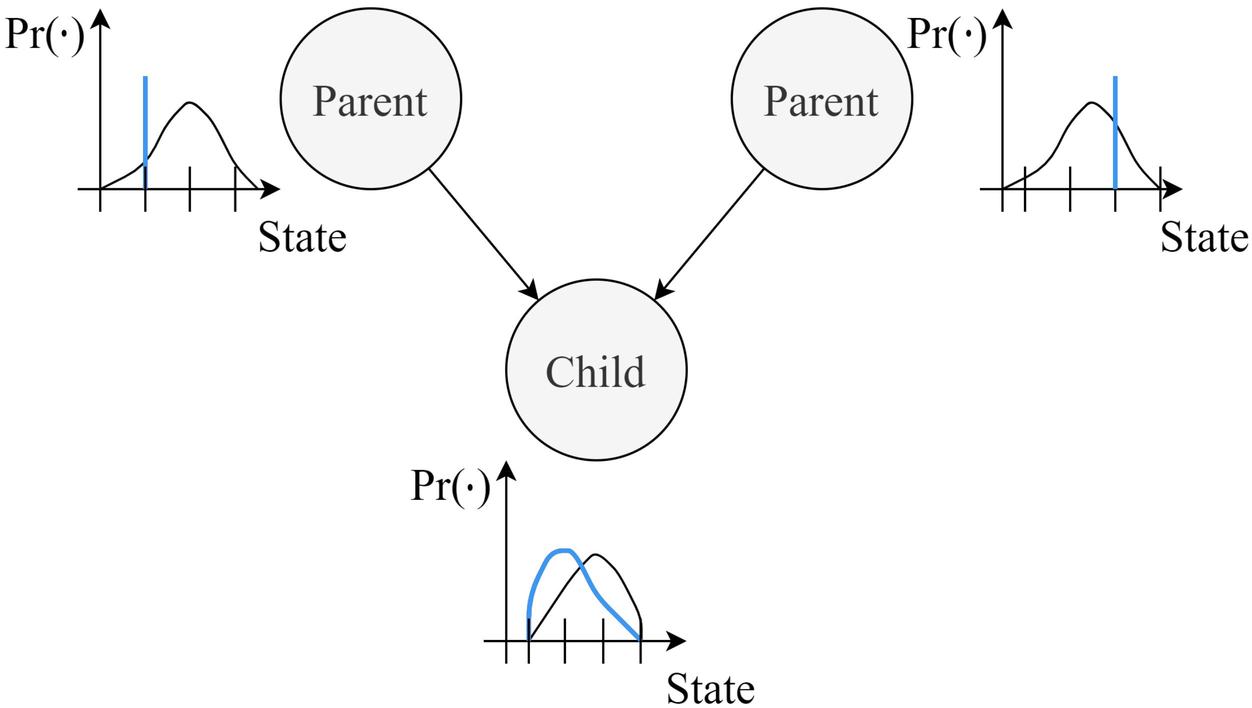


Figure 4.9: Construction idiom between two parent nodes and a child node. Note that this dependency may seem similar to the Causal idiom, however the meaning of the relationship is distinct. The parent nodes compose (with uncertainty) the child node. This relationship may be consolidated into a single node, although the uncertainty will not be preserved.

into a single node, at the cost of masking the uncertainty in the relationship.

In the Example Scenario, consider the PIF Workload that could be decomposed into sub-PIFs such as “Task Load” (for goal-relevant tasks) and “Non-Task Load” (for necessary, but not goal-relevant tasks). While the extreme states of Workload may be clear (e.g., Workload is High when both Task Load and Non-Task Load are High), the intermediate states are uncertain. For example, assume that the crew has a low number of goal-relevant tasks but many administrative tasks (required, but not goal-relevant). It could be reasonably argued that Workload is High, Medium, or Low, depending on the analyst’s perception of the scenario. As a result of this uncertainty, a definitional dependency should not be assigned in this scenario. Nor should a causal dependency be assigned, because the relationship is definitional/decompositional rather than causal. Quantification of construction dependency may leverage, for example, fuzzy sets and logic as a method for allowing analysts to uncertainly prescribe the definition of the child node on the basis of the two parents.

4.3.4 Common Context

Common Context dependency describes a relationship in which two (or more) children nodes share at least one common parent node. This characterizes the common-cause mechanisms described by NUREG/CR-6265 ([24]), and is similar in concept to the indirect dependence in THERP ([16]). The child nodes are dependent on each other, but not directly causally related; the causal relationship stems from the common parent(s). Although there are at least three nodes in this relationship, the dependency is described as between the child nodes sharing a common parent. In a Common Context dependency, the parent node(s) are the drivers of the dependency, but it is important to recognize that the child nodes are therefore dependent on each other.

This relationship is likely common between the PIFs, which represent an inherently and intractably interdependent set of factors, as well as between the PIFs and CFMs. This dissertation defines "context" as the set of PIFs describing static conditions that are invariant during the portion of scenario under study. Context PIFs describe organizational, environmental, and/or process-related factors that will not change due to events during the time of interest. For example, an organization's safety culture or work processes are determined externally to any scenario and will not change through the course of a scenario or time of interest. Thus, they represent a *common context* shared by the CFMs relevant to different MCFs, and thus the MCFs are linked via Common Context dependency. The dependency factors used in, e.g., THERP, SPAR-H, and Fire HRA (closeness in time, crew similarity, procedures, etc.) actually identify common contextual features between the tasks, although without acknowledging the causal pathways. Therefore, Common Context dependency characterizes the dependencies typically included in THERP-derived methods, and provides the causal context of these relationships[10, 16, 31]. Figure 4.10 depicts both a generic Common Context idiom between two child nodes, and a Common Context dependency between two CFMs driven by a parent set of PIFs (the *context*), which is expected to be the most common instantiation of the idiom.

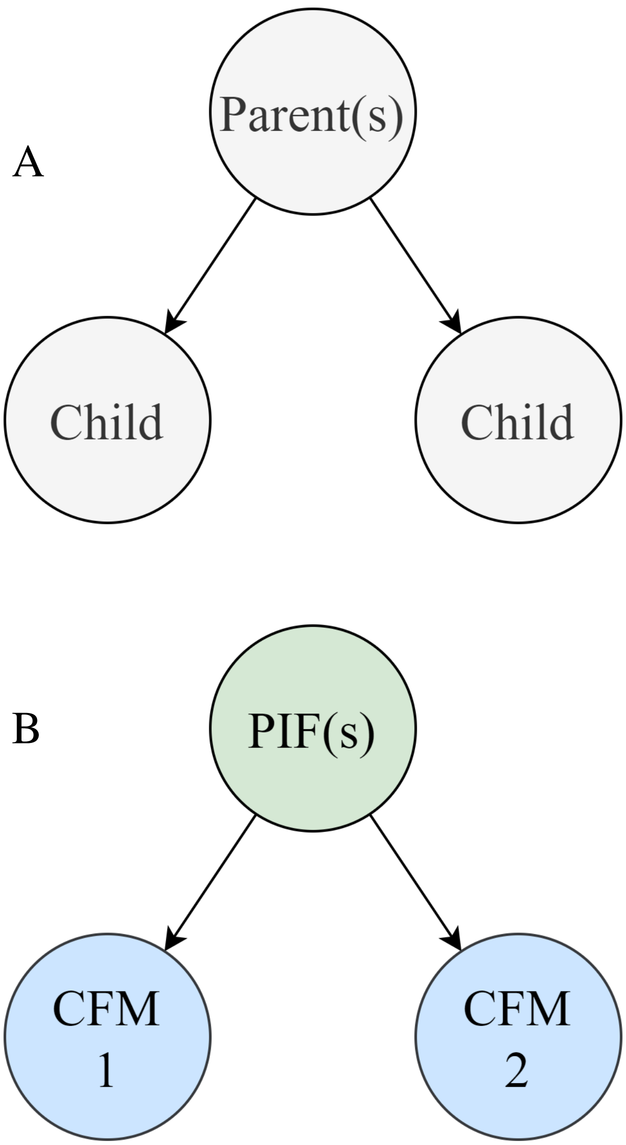


Figure 4.10: (A) Generic Common Context idiom between two child nodes sharing a common cause parent node. (B) Common Context idiom between two CFMs which share a parent PIF node (the context).

In the Example Scenario, both MCF A and MCF B are performed in the main control room by the same crew. While Workload, Time Available, and Stressors will vary between MCF A and MCF B, Training/Experience and Fitness for Duty do not change between MCFs. The crew does not increase or decrease their training/experience level or their fitness for duty between MCF A and MCF B. Conversely, the workload, available time, and stress level will fluctuate throughout the scenario. Therefore, both MCF A and MCF B share the common set of PIFs {Training/Experience, Fitness for Duty}, and will be dependent on each other as a result of the common context.

4.3.5 Situational

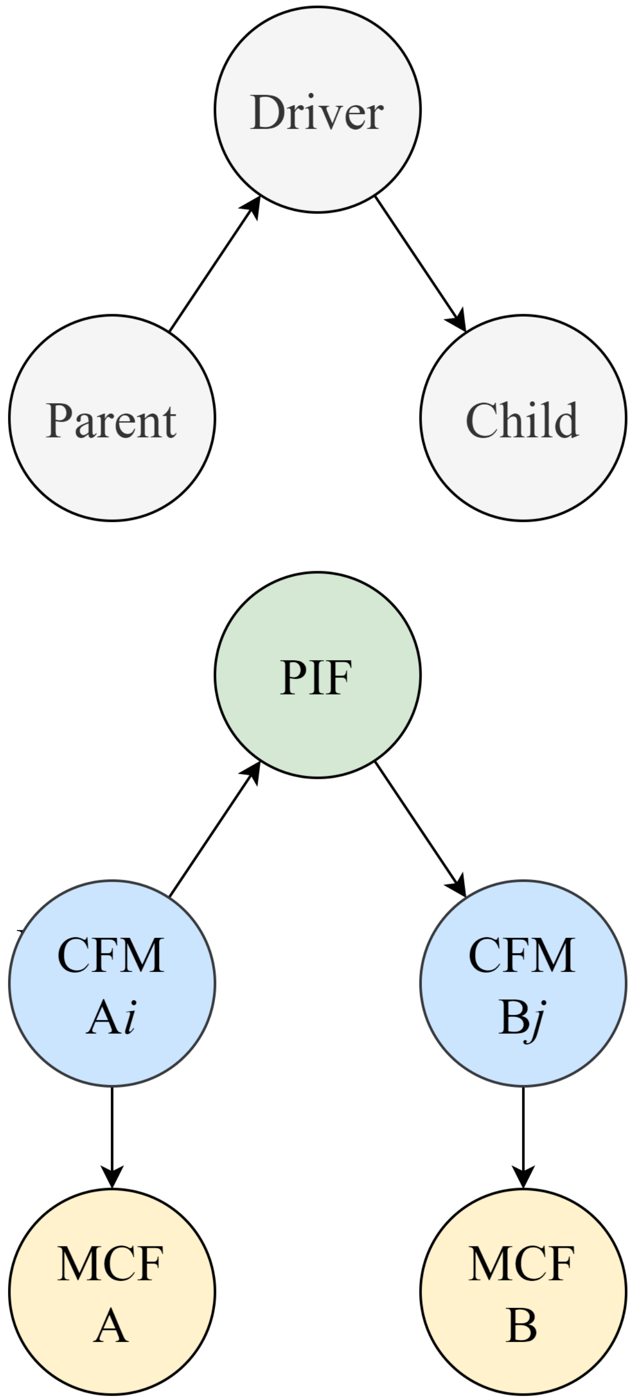


Figure 4.11: (A) Generic Situational idiom with a single parent, driver and child node. (B) Situational idiom between two CFMs where *CFM 1* drives the PIF(s) (the situation) surrounding *CFM 2*.

Situational dependency describes a causal chain relationship that exists across time, i.e., acts between MCFs. An HRA construct instantiated at one time (i.e., during one MCF) causes a change in a future HRA construct. For instance, a failure of Task A causes a change in the *situation* (i.e., local PIF set) surrounding Task B. Task B. the third node in the causal chain, thus experiences a change in the probability of failure as a result of its occurrence within the new situation dictated by the causal chain from Task A. The Situational dependency idiom characterizes the Error-Forcing Contexts from ATHEANA ([48]) and NUREG/CR-6265 ([24]). Situational dependency is similar to Common Context dependency in using a third variable to drive dependency between two other variables. However, instead of the dependency driver being a common (static) context, Situational dependency is driven by an intermediary dynamic PIF that transmits the relationship across time. This dissertation defines situation as distinct from context in its dynamic nature: whereas context is invariant in the scenario, the situation varies as a result of previous failures and PIF changes.

Figure 4.12 visualizes the difference between the *context*, which describes the overarching and static set of scenario characteristics, and the *situation*, which describes the variable and MCF-specific set of task characteristics. The PIFs surrounding MCF A and MCF B may be the same constructs, but change with time and previous performances (via situational dependency). Context PIFs, on the other hand, do not change with time either as a natural variation or the result of previous performances.

Situational dependency describes the changes engendered in PIFs around a task/function as the result of previous failures; the altered PIF states in turn change the probability of MCF failure. Situational dependency is therefore an indirect dependency between CFMs for different MCFs, because the relationship is passed through a third variable (the PIFs). However, it is a direct dependency in the sense that there is a continuous causal chain between the parent node(s) and child node(s). Situational dependency is a causal chain structure: the directed arcs defining the relationships are oriented in the same direction, from the parent node through the driver node to the child node. In the Example Scenario, the failure of MCF A may change the PIFs surrounding MCF B. Further, specific CFMs of MCF A will affect the PIFs differently. For instance, a Delayed Diagnosis on MCF A will reduce the Time Available and may increase the Stressors for MCF B. However,

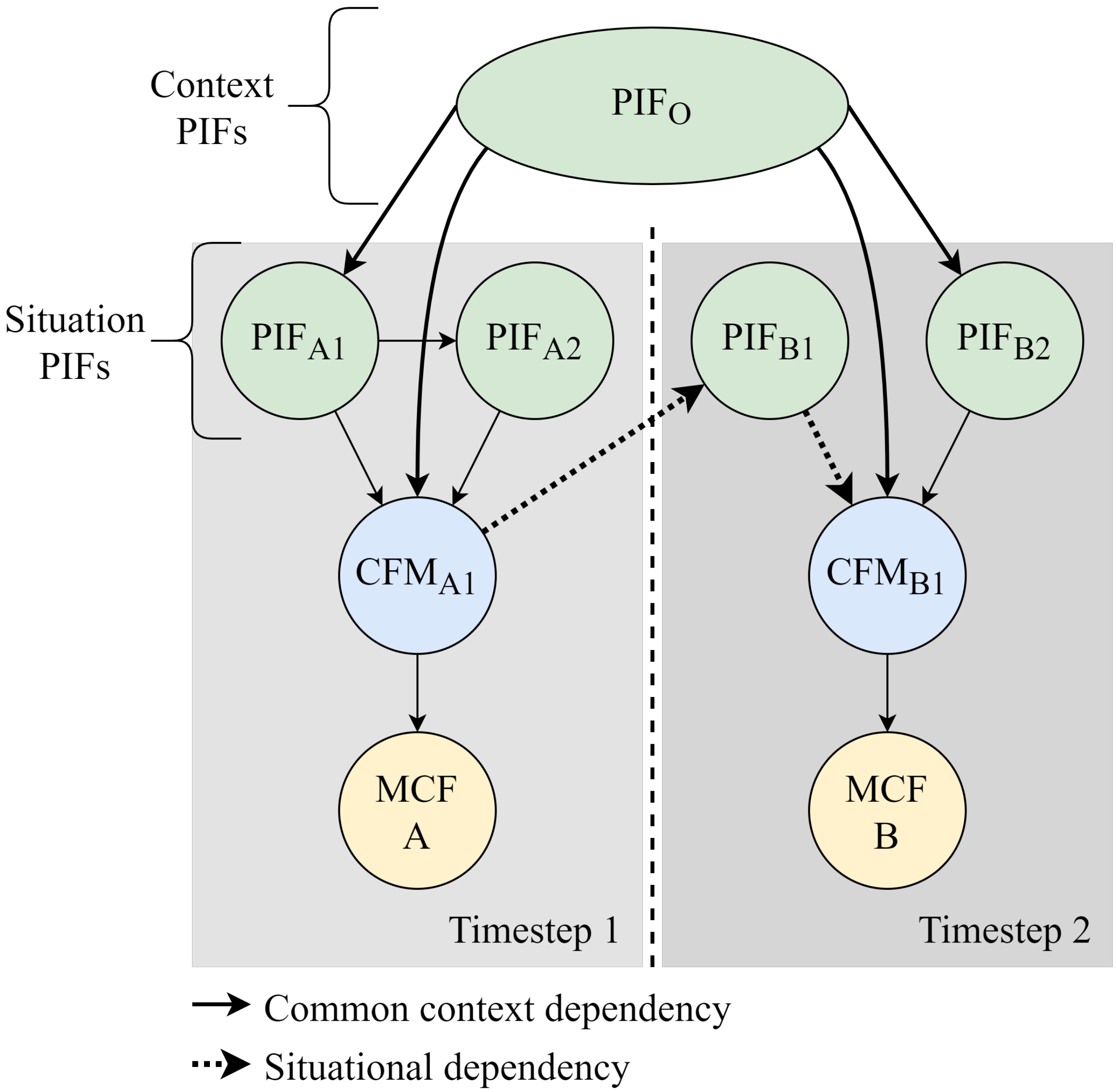


Figure 4.12: Context and situation as visualized on a Bayesian network.

Incorrect Diagnosis may increase the Workload during MCF B as the crew needs to recover the failure once discovered. Different CFMs will therefore affect the situation for subsequent MCFs in unique ways, and so the CFMs (when available in the model) should generate the situational dependencies.

Situational dependencies are expected to extend from the CFM(s) of a prior MCF to the CFM(s) of a future MCF, as shown in Figure 4.11. This relationship therefore can connect MCFs to each other, and describes relationships that previously might have been modeled as direct task-task or HFE-HFE dependencies. Situational dependency replaces the "error begets error" concept with the realistic causal structures that describe the actual relationships taking place, and allows for modeling more complex relationships. Because MCF and HFE nodes are uninformative in the BN model, the CFMs propagate situational dependencies to influence future PIFs and CFMs (and, therefore, future CFMs and HFEs).

4.3.6 Effect-modulating

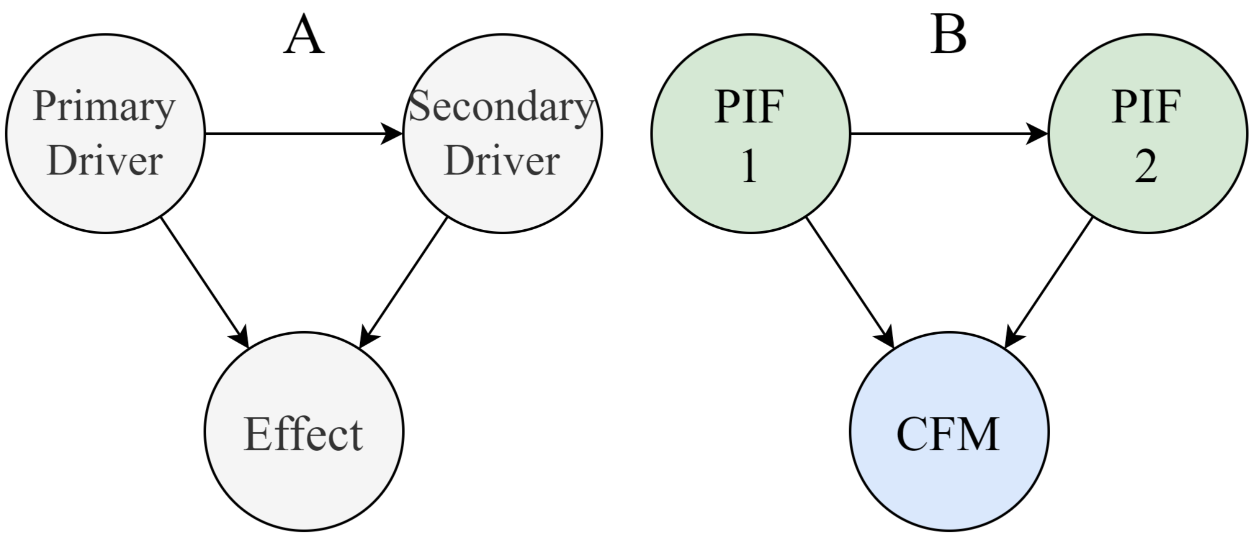


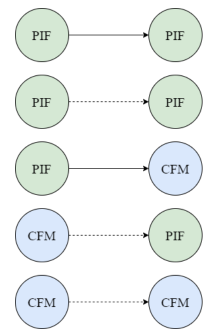
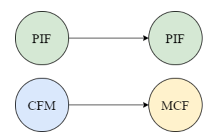
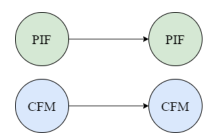
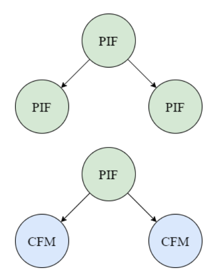
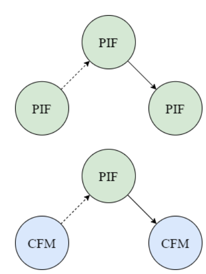
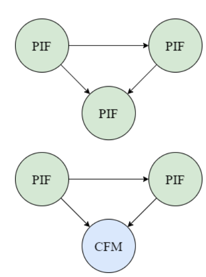
Figure 4.13: (A) Effect-modulating dependency where the non-linear effects of PIF 1 and PIF 2 are stored in the node PIF 3. (B) Effect-modulating dependency where the non-linear effects of PIF 1 and PIF 2 are included directly in the calculation of the states of the CFM node.

Effect-modulating dependency captures the non-linear effects of multiple HRA constructs acting together. This is an important aspect of HRA that may be particularly characteristic of PIF inter-relationships [35]. Quantifying the effects of multiple PIFs is a familiar problem for HRA; for example, SPAR-H incorporates a correction factor when multiple (≥ 3) negative PIFs are present. However, this does not reflect the effects of the PIFs but rather curtails the possibility of recording an invalid HEP (i.e., HEP > 1) given multiple negative PIFs in SPAR-H [31]. Effect-modulating dependencies, on the other hand, exactly reflect these non-linear effects between PIFs; the effect can be modeled by either a “macro-PIF” or directly into the relevant CFM, as shown in Figure 4.13.

Consider again the Example Scenario, and specifically the interrelationships between the PIFs Time Available, Workload, and Stressors. In this case, Workload can be considered inflexible: the workload is assumed to be an artifact of the scenario that does not vary based on how the crew operates. However, both Time Available and Stressors will be affected by crew actions. Workload affects both Time Available and Stressors: a high workload may reduce the time available for goal-relevant tasks, and may increase the stress on the crew. Low time available can also increase the stress on the crew. A high workload *and* low time available may *exacerbate* the stress on the crew, and result in stress higher than either high workload or low time available individually. Conversely, high time available and low workload may *mediate* the stress on the crew, and result in stress *lower* than either low workload or low time available would produce individually.

Effect-modulating dependencies consist of a “primary driver,” “secondary driver,” and “effect” construct. Delineating between the three may be non-trivial, but revelation is facilitated by considering pair-wise causal relationships in the acyclic constraints of the BN architecture. The primary driver is a (probabilistic) cause of both the secondary driver and effect. The non-linear effects of PIFs can manifest in two distinct manners: exacerbating or mediating. These descriptors are in relation to the ultimate/resultant propensity toward failure. If the component PIFs exacerbate each other, the resultant state will be worse than the component PIF states would sum to (resultant is “greater than the sum of its parts”). In this context, “worse” conveys a higher propensity to causing human performance issues. If the component PIFs mediate each other, then the resultant state will be better than the component PIF states would sum to (resultant is “less than the sum of its parts”). As Figure 4.13 shows, the resultant state may be stored in a separate PIF, or may be incorporated implicitly in the computation of CFM state probabilities.

Table 4.1: HRA dependency idioms and visualizations. Dashed arrows indicate relationships across MCFs.

Idiom	Definition	HRA Examples
Causal	A relationship describing a causal connection between two or more elements. A change in the parent node (cause) state inflicts a change in the probability of the child node (effect) state.	
Definitional	A relationship describing a definitional or synthetic connection between two or more elements. The parent node(s) comprise the child node. A change in the state of the parent node inflicts a deterministic change in the state of the child node.	
Construction	A relationship describing an uncertain definitional relationship between two or more elements. The relationship appears both causal and definitional, and node consolidation is applicable. A change in the state of a parent node inflicts a change in the probability of the child node state.	
Common Context	A relationship describing a common-cause relationship between two elements sharing a (set of) parent node(s). This relationship is a combination of causal idioms connecting the parent node(s) to each child node. A change in the state of the parent node inflicts a change in the probability of each child node state. A change in any child node state inflicts a change in the probability of the parent node state(s) and the other child node states.	
Situational	A relationship describing a serial connection of multiple elements, at least one of which is an intermediate PIF. This relationship is a combination of causal idioms connecting each node in the series. A change in the state of any node inflicts a change in the probability of the other node states. This is realized in HRA as a CFM or PIF influencing the situation (PIF(s)) affecting other CFMs or PIFs.	
Effect-modulating	A relationship describing the non-linear effects of multiple PIFs interacting with each other. PIFs may exacerbate or mediate each other, resulting in an aggregate PIF state which is worse or better than a linear combination of the PIFs would suggest.	

4.4 Result: HRA Dependency Idioms for Quantitative HRA

4.4.1 HRA Dependency Idioms

The HRA dependency idioms developed in the previous section describe distinct causal structures that exist between HRA constructs and describe different dependencies. The idioms are distinct and discrete structures that characterize how HRA constructs are causally and definitionally aligned, and therefore structure the dependency relationships between them. As discussed in Section 4.3, six HRA dependency idioms were created on the basis of previous idiom structures (from [43, 100]), HRA constructs defined in Chapter 3, and the semi-hierarchical HRA-BN structure created in this chapter (described in [30, 44, 88]). The six idioms are defined and visualized in Table 4.1.

Briefly, the six HRA dependency idioms are defined:

- **Causal**: A probabilistic cause-effect relationship between two HRA constructs where the occurrence of the cause changes the probability of the effect occurring.
- **Definitional**: A deterministic constituent-whole relationship between two HRA constructs where the occurrence of the constituent precisely determines the state of the whole.
- **Construction**: A probabilistic constituent-whole relationship between two HRA constructs where the occurrence of the constituent changes the probability of the state of the whole.
- **Common Context**: A probabilistic common-cause relationship between three (or more) HRA constructs. The occurrence of one effect changes the probability of the other effect.
- **Situational**: A probabilistic causal chain relationship between three (or more) HRA constructs. The occurrence of any construct in the chain changes the probability of the others.
- **Effect-modulating**: A probabilistic exacerbating or moderating relationship between three (or more) HRA constructs. Certain state combinations of the two drivers will result in a more (exacerbating) or less (moderating) error-prone condition than either would alone.

4.4.2 HRA dependency Idiom Instances in Data

The idioms represent the different types of dependency that are possible given the HRA constructs and semi-hierarchical BN architecture previously defined. Each idiom is a distinct logical structure associated with a BN representation, and is applicable to multiple HRA constructs. As a result, there will be several *instances* of each idiom observable in HRA data. While Table 4.1 shows applications of the idioms between different HRA constructs, an *instance* is even more specific. An idiom instance is an observation of a dependency idiom involving particular HRA constructs (i.e., specific PIFs, CFMs, and/or MCFs). Dependency idiom instances can be identified in HRA data. This section presents idiom instances identified in HRA data, principally in scenarios from NRC's SACADA database [84]. It should be noted that SACADA does not record failure modes. Therefore, MCFs are used in place of CFMs for situational dependency idioms, essentially treating the UNSAT state of the MCF as a CFM. The definitional dependency between CFMs and MCFs allows CFMs to be represented as states of the MCF, so this replacement is acceptable.

Causal dependencies are expected to be present between many HRA constructs. SACADA defines the Situational Factors (SFs, equivalent to PIFs) as *causal* factors relevant to operator performance [84]. Therefore, causal dependency exists between the relevant SFs and each Training Objective Element (TOE, equivalent to MCF). Definitional dependencies are expected to be constrained mostly to CFM-MCF relationships, because CFMs define the failure of the MCF (i.e., an MCF can either be successful or failed via one of its CFMs, and therefore its states are Success; CFM 1; CFM 2; etc.).

Construction dependencies will be present in PIFs that can be uncertainly decomposed into more fundamental components. The PIF taxonomy defined in [42] incorporates this expanding decomposition structure. In this structure, however, it is uncertain how lower-level constructs combine to form the higher-level PIFs. For example, "Training Program" can be decomposed into lower-level constituents "Availability" and "Quality," but there is uncertainty in the relationship. If the lower-level constituents are low availability and high quality, the higher-level PIF Training Program state is not precisely determined. Thus, this necessitates a probabilistic relationship (that could be quantified with, e.g., fuzzy sets). In SACADA, the "Overarching Issues" PIF has disparate states (that could stand alone) that are combined into a single factor, but the manner in which the states form a single coherent entity is unclear [84].

Common context dependencies will be present between CFMs for MCFs that occur within the same invariant context. Such PIFs, external to the dynamics of the scenario, will not change across the time of interest and therefore impose a Common Context dependency among affected MCFs. For example, the first two TOEs in a Loss of Feedwater (LOFW) scenario are both performed in a noisy background, and therefore are dependent via their common context. In other scenarios, TOEs are performed in common contexts related to multiple external demands and non-standard event sequences [84]. SACADA includes a set of Overarching Factors PIFs that may serve as a common context for performance and imposing dependencies between many MCFs, as Chapters 5 and 6 will show.

Situational dependencies are expected between CFMs in sequential MCFs, wherein the occurrence of an initial CFM changes the situation (i.e., the PIFs) surrounding the second MCF. For instance, failing to perform a procedural task can lead to low available time to perform subsequent tasks, as occurred in SACADA. A crew failed to navigate the procedure to stop pumps, and accordingly had barely adequate time to perform the next action [84].

Effect-modulating dependencies will be present between PIFs that interact together to produce nonlinear positive or negative effects. While this is more difficult to identify at first-glance in data, preliminary results reveal that such relationships are plausible between, e.g., Time Available and Workload PIFs. In this instance, the Workload (number of concurrent tasks) during a given MCF will affect both the available time to complete the MCF *and* the probability of failure on the MCF itself. The result is that certain state combinations (e.g., Multiple Concurrent Tasks and Barely Adequate Time) result in much greater probability of failure (than the marginal), while others (e.g., Nominal Workload and Expansive Time) result in much lower probability of failure (than the marginal value).

4.5 Discussion

The set of HRA dependency idioms created in Section 4.3 provides the architecture to conceptualize and traceably model complex dependency relationships in HRA. This chapter implements the causal, probabilistic definition of dependency presented in Section 3.6 and provides distinct causal structures for the relationships possible between the HRA constructs defined in Chapter 3 and the necessary relationships identified in [30]. The idioms for HRA translate the established causal logic structures defined in [43, 100] to improve the foundational knowledge of dependency for HRA. The HRA constructs and necessary relationships from [30] show that there is a finite set of dependency relationships that must be included in an HRA model. Further structuring the HRA constructs within a BN (i.e., Figure 4.1) imposes additional restrictions on the possible dependency relationships. The logic structure of the BN means that some dependencies (e.g., PIF-HFE and other multiple-level relationships) can be immediately removed from the model, as HRA constructs will only interact at their "level" and/or with constructs in adjacent levels of the hierarchy. Similarly, defining the CFMs as mutually exclusive means that there are no inter-CFM causal dependencies [44]. The idioms were created to both add formality and causal structure to existing dependency concepts (e.g., the direct and common-cause dependency relationships) as well as characterize the more complex relationships that occur within and between MCFs.

4.5.1 Idioms Relationship to Foundational Gaps in Dependency

The HRA dependency idioms are not a wholly new concept, but leverage and adapt existing constructs in BNs ([43, 100]) to the HRA context. The idioms are not limited to positive (i.e., "error begets error") dependency, and represent the multiple types of dependencies that exist between different HRA constructs (rather than only between tasks or HFEs). The idioms therefore address the limitations imposed by the "error begets error" paradigm of previous dependency frameworks, and move beyond addressing dependency only between tasks or HFEs. Further, the idioms facilitate the direct quantification of conditional probabilities from HRA data. Thus, the idioms facilitate developing data-driven quantification methods to replace previous, unfounded quantification scheme(s) for dependency. Finally, the idioms characterize the causal basis underlying dependency relationships from other frameworks. Thus, the idiom framework addresses many of the foundational gaps associated with HRA dependency.

Beyond addressing the foundational gaps in HRA dependency, the idioms also provide traceability to the process of identifying and assessing dependency by using repeatable graphical structures to model dependency relationships. The objectivity of dependency assessment is similarly maintained by visualizing dependency and thus facilitating peer review of the assessed dependencies in any scenario. The idioms will improve the objectivity and accuracy of dependency assessment by supporting the quantification of conditional probabilities from HRA data, with minimal reliance on expert elicitation and/or anchoring values. Finally, the idioms enforce a formative, rather than summative ([15]) process for analyzing dependency; the BN model for a scenario is fully constructed with included dependencies *before* quantification. This ensures that the structure of the model and qualitative aspects of HRA are attended to appropriately, and ensures that no part of the process overrides any other.

The HRA dependency idioms created herein represent fundamental dependency relationships that can exist between HRA constructs. As indicated above, the idioms address many of the foundational gaps in HRA dependency identified in Chapter 2.

4.5.2 Impact of the HRA Dependency Idioms

The idiom framework for HRA dependency is a novel way of conceptualizing and modeling dependency in HRA with a basis in probabilistic causality. They capture dependencies used in previous dependency frameworks with an expanded causal basis, and describe more complex dependencies that are possible between HRA constructs. As a result, the HRA dependency idioms are expected to have both short-term and long-term impacts on the process of HRA dependency analysis.

4.5.2.1 Short-term Impact

In the short-term, the HRA dependency idioms can provide more traceability and objectivity to existing dependency assessment frameworks. The idioms enforce causality as the basis for the relationships that are currently considered in dependency analysis (that fall under, e.g., common context and situational idioms). Thus, current dependency assessments can be made more traceable by referring to the HRA dependency idioms as known relationship structures and using causality as the basis for dependencies rather than the checklists of factors in common.

4.5.2.2 Long-term Impact

In the long-term, the HRA dependency idioms form the basis of a more comprehensive framework that emphasizes formative, causal, data-based dependency analysis. The idioms can be used to decompose and verify existing dependency assessments, but perhaps more importantly they provide the building blocks for the initial constructing of BNs for HRA. This is *formative* dependency, built into the model from the beginning, in contrast to the current *summative* model where dependency is applied at the end of the quantification process [15]. This ensures that dependency analysis is considered in a traceable manner and does not override other steps of the HRA process. Within this formative process, the HRA dependency idioms ensure that assigned dependencies are rooted in known mechanisms that operate between the HRA constructs. For definitional dependency, the basis is the mathematical formalism of the set-subset relationship. For other dependencies based on causality, the basis is the real (physical, phenomenological, and/or cognitive) mechanisms that connect HRA constructs. Using the robust bases of dependency ensures that dependency relationships are well-founded and assigned in a logical manner. Finally, the HRA dependency idioms facilitate the direct quantification of dependency from HRA data, as will be demonstrated in Chapters 5 and 6. This means that expert elicitation and arbitrary multipliers will be replaced by data-backed conditional probability assessments, further strengthening the technical basis of HRA.

Chapters 5 and 6 demonstrate this impact of the HRA dependency idioms by applying them to a case study from HRA data. The idioms are used to build BN submodels for each MCF and a full BN of the scenario in Chapter 5, and guide the quantification of the network in Chapter 6. The case study therefore serves as a proof-of-principle for the long-term impact of the HRA dependency idioms.

4.6 Conclusions

The HRA dependency idioms presented in this chapter implement the robust definition of HRA dependency rooted in causality ([44]), and capture the spectrum of dependency relationships that exist between HRA constructs. The HRA dependency idioms are graphical manifestations of dependency as a probabilistic, causal relationship between two or more HRA constructs. The HRA dependency idioms form a set of basic relationships for HRA that guide analysts immediately in identifying dependencies between HRA variables for incorporation into current HRA models. At a high level, the idioms provide enhanced traceability and objectivity to the current techniques for analyzing HRA dependencies. Relationships solicited for use in current HRA methodologies can be traced to specific causal structures (idioms) rather than a set of vaguely-defined factors to assess a single relationship type. The HRA dependency idioms will also facilitate the wider use of causal BNs for HRA. The idioms simplify the construction of complex HRA BNs by providing distinct structures to identify relationships between a manageable number of variables, which can then be connected into larger BNs to form an entire HRA model. Building the network from the ground-up in this manner makes HRA model building easier, and the use of idioms improves the traceability of the final HRA model. The idioms can also be used to decompose previously-built HRA BNs to identify and validate the encoded causal dependencies.

The HRA dependency idioms presented here may not be an exhaustive set of the distinct relationships possible in HRA, although it is posited that this set of idioms is comprehensive. In the case further HRA dependency idioms exist, the foundations of this work (i.e., the delineation of distinct logic structures) provide the ability to identify new idioms. The HRA dependency idioms provide the qualitative meaning and structure underlying different relationships, and facilitate quantification of dependency with HRA data. The quantification of dependency relationships is developed in Chapter 6.

Chapter 5: Building BNs from HRA Data Using the Idiom Structure

This chapter addresses the first portion of Research Objective 3, the development of the mathematical framework for quantifying dependency (both the idioms and full-scope BN constructions). This chapter creates the process for building the BN structure from the HRA dependency idioms, available HRA data, and expert causal knowledge, thereby connecting the causal reasoning structures (HRA dependency idioms) developed in Chapter 4 to the quantification of full HRA BN models.

5.1 Introduction

The HRA dependency idioms created in Chapter 4 provide the capability to understand and model multiple types of dependency between the HRA constructs. These therefore provide a significantly improved qualitative understanding of the nature of dependency. However, understanding why, how, and how often human error occurs in complex engineering systems requires building robust causal models based on the HRA dependency idioms to rigorously model and quantify the dependencies observed in an HRA scenario. Doing so necessitates a structured methodology for building a full model from the HRA modeling constructs/variables and the HRA dependency idioms. Further, robust models must be traceable and objective. This chapter develops the methodology for traceably and objectively construct HRA causal models from the HRA dependency idioms.

Accordingly, this chapter reviews two available HRA data sources and presents the case study scenario that was used to develop the BN construction methodology based around the HRA dependency idioms. Then, this chapter details the processes involved with cleaning HRA data and mapping the data to modeling constructs, as well as the procedures available to down-select or prune the model, to ensure robustness while limiting the computational resources required to build and quantify the model. This chapter then details the methodology for creating HRA BN models using the HRA dependency idioms as the framework for building the causal structure of the model. The methodology created in this chapter is a traceable and objective way to create HRA BNs including formative dependency, and thus is a significant improvement for HRA modeling.

5.2 Approach: HRA Data Source Selection

The first consideration for building HRA Bayesian network models from data is the selection of the data source(s) that will be used to define the variables (nodes) used in the model and populate the conditional probability tables (CPTs). In this context, "data source" is a broad term that encompasses the typical connotations of data (e.g., data tables, simulations, and experiment results), as well as prior knowledge (e.g., expert elicitation and assumptions) and narrative/qualitative items (e.g., crew interviews, observation recordings, and variable names). All of these are sources of data that can, and should, be incorporated for use in building and quantifying HRA BNs. For the purposes of this study, however, the focus will be on using simulator/experiment results and narrative/qualitative elements, rather than expert elicitation and assumption-based elements, that can be used to build HRA BNs. In light of that focus, there are two main HRA data sources available that report the results of simulator/experiment studies: the U.S. NRC's Scenario Authoring, Characterization, and Debriefing Application (SACADA) ([84]) and the Human Reliability Data Extraction (HuREX) framework from the Korea Atomic Energy Research Institute (KAERI) [102]. These methods will be briefly reviewed in the following subsections in regards to the qualitative/narrative and quantitative data supported by each. For a thorough comparison of both methods, readers are referred to [103, 104]. This study used the NRC's SACADA database as the sole source of HRA data.

5.2.1 HRA Data Source: HuREX

The Human Reliability Data Extraction (HuREX) framework was developed by the Korea Atomic Energy Research Institute (KAERI) to provide a standard methodology for obtaining HRA data (principally human error probabilities (HEPs), but also correlations between HEP and PIF state values) from operator crews as they perform tasks in full-scope nuclear reactor simulators. HuREX has also been used to collect data from the observations of real control rooms [102]. HuREX incorporates the known/planned scenario information, operations logs, audio/video recordings and crew interviews to collect data predominately at the subtask and/or task primitive level, which corresponds roughly to the crew activity primitive (CAP) defined in Section 3.4.1 [103]. The HuREX method also identifies Unsafe Acts (UAs)[1], which are analogous to HFEs, and relates their occurrences to specific procedural steps, tasks, and cognitive activities to identify possible failure pathways. HuREX identifies 89 generic data items (GDIs) about the environment, human-machine interface (HMI), organizational factors, procedures, tasks, success criteria, and the actual performance context. The GDIs guide the investigators to variables that should be investigated as PIFs for a scenario, but not all 89 GDIs are incorporated for each task [103].

The nominal products of HuREX studies are updated estimates for the HEPs of various task types and evaluations of the correlation between HFEs/UAs and associated PIFs. More importantly, however, HuREX has provided another source of data in a data-scarce field [102]. Additionally, studies using HuREX have produced valuable empirical probabilities ([106, 107]), insights into the veracity of HRA assumptions regarding dependency ([98, 108]), and guidelines for effective data collection ([109]). As of 2022, HuREX has collected 45,000 data points from 223 simulator studies, making it a valuable data source for HRA researchers [103].

The HuREX framework is not without its challenges. Implementing HuREX for a study requires a great deal of front-end and back-end effort from the observers and analysts. Unlike other data collection schemes (i.e., SACADA), HuREX requires analysts to define the relevant PIFs as a subset of the 89 GDIs [102]. Furthermore, the level of task detail at which HuREX operates (i.e.,

[1] Other conceptualizations of the Unsafe Act are more similar to "near misses" or "pre-errors" than HFEs [105].

the task primitive or CAP) may not be optimal for some HRA use cases. This challenge may be overcome by synthesizing higher-level tasks or functions (i.e., MCFs) from the HuREX data using construction dependency to create the modeling constructs, although this may induce additional biases and uncertainty.

5.2.2 HRA Data Source: SACADA

The Scenario Authoring, Characterization, and Debriefing Application (SACADA) framework was developed by the U.S. NRC to provide a usable methodology for collecting and analyzing data from full-scope simulator studies, as well as predicting human reliability in real scenarios (i.e., scenarios that operators may actually face in a main control room) [84]. Like HuREX, SACADA is a worksheet-based approach in which analysts/observers identify the assigned or observed PIF states during each step of the simulation. SACADA collects over 40 distinct Situational Factors (i.e., PIFs), most with multinomial (i.e., more than 2) states. Available SACADA data comes from full-scope simulator studies. SACADA has recorded over 1,300 simulated scenarios since its creation ([103]), encompassing over 15,000 recorded tasks/functions. SACADA collects data via observations and crew interviews; observational data assigns states to the PIFs for each Training Objective Element (TOE), which are the MCFs performed for each given scenario. SACADA data, therefore, consists of scenarios decomposed into sequential functions, each encoded with the states of relevant PIFs and an overall parity value ("UNSAT" or "SATΔ" for failures, "SAT" or "SAT+" for successes)[2]. For consistency of terminology with the rest of this manuscript, "MCF" will be used to refer to SACADA TOEs and "PIF" will be used instead of SACADA Situational Factors.

SACADA data is split into two main datasets: the set of all MCFs performed in a scenario (the main dataset), and the set of only failed (UNSAT) MCFs recorded in each scenario (the failed events dataset). There are 29 PIFs available to encode the main dataset, with an additional 23 PIFs applied to the failed events dataset [84]. The encoded data is stored in a spreadsheet, an example of which is provided in Appendix A.

[2]SACADA does not qualify SATΔ as failure, but this research includes them as they represent deficient operations by the crew

Using the SACADA database for this work imposed some challenges, mainly due to the sparsity of the data recorded. HRA is a data-sparse field in two senses: the lack of available data generally, and the lack of events (i.e., recorded failures) in the data that is available. The latter form of sparsity is also referred to as "low-Event per Variable" or low-EPV data, in which there is a significant discrepancy between the number of variables (PIFs and their states) recorded and the number of events (failures) observed. The presence of low-EPV data has a heavy influence on model choice for predictive modeling applications, as will be discussed in Chapter 6. Bayesian networks greatly improve the usability of low-EPV data, because it can be augmented with engineering knowledge and expert elicitation, but this is still a concern for using SACADA data.

5.3 Approach: Case Study Scenario Description

The methodologies presented in this chapter and Chapter 6 were developed and proven through the use of a case study on a real scenario in the SACADA dataset. The scenario was a simulated inter-system loss of coolant accident (ISLOCA) that consisted of 13 distinct MCFs that the crew(s) needed to complete. The network produced for this case study is anonymized, so that each MCF is represented only as an IDAC cognitive phase instead of a specific task. This also allows the HRA dependency idioms to be re-used in models of other scenarios and generalizes the quantification from individual MCFs to cognitive phases.

Each cognitive phase that appears in the model is represented as an identical BN submodel (i.e., each D-phase submodel is identical) consisting of the MCF with its associated PIFs. PIFs were assigned to each cognitive phase MCF from relevant SACADA SFs using the processes described in Section 5.5. The following section describes the process of mapping the SACADA data to create the modeling constructs for this BN. The case study for this research consisted of the following sequence of cognitive phase MCFs:

1. Decision & Diagnosis
2. Action Taking
3. Decision & Diagnosis
4. Decision & Diagnosis
5. Communication
6. Information Gathering

7. Decision & Diagnosis
8. Action Taking
9. Action Taking
10. Decision & Diagnosis
11. Decision & Diagnosis
12. Action Taking
13. Decision & Diagnosis

5.4 Approach: Mapping HRA Data to BN Structures

The HRA data source used in this research, the U.S. NRC's SACADA dataset, required some preprocessing in order to facilitate building the BN model. For instance, data cleaning and sorting was required to develop the network structure and prepare for the creation of the CPTs. This dissertation provides methodologies for building and quantifying BNs using the HRA dependency idioms using SACADA constructs directly. The same processes could be done using existing HRA methods (e.g., SPAR-H, Phoenix, or IDHEAS), although additional translation or data preprocessing may be required. For example, using SPAR-H may require translating and/or combining the SACADA PIFs (i.e., Situational Factors) into the more constrained set of 8 PIFs used in SPAR-H.

The methodology developed in this chapter is also amenable to using expert-driven processes for model construction in addition to, or instead of, available HRA data. Existing HRA methods may also be used, although some constructs used in the methods may not be present in the available data (e.g., CFMs are used in Phoenix and IDHEAS, but are not explicitly found in SACADA). When using expert-elicited modeling constructs instead of the constructs contained in a database, the data cleaning and mapping processes discussed in this section will not be required. This section deals specifically with translating HRA data to build the nodes in a BN. Although the data used in this research was obtained from SACADA, the processes are described generally to facilitate the use of other existing or novel data sources as available.

The case study model that was used to develop this methodology was built with SACADA data and exert knowledge. The Situational Factors were incorporated as PIFs, and the TOEs became the MCFs. The TOEs in SACADA are high-level constructs (as opposed to the lower level elements in, e.g., HuREX), making them more appropriate for MCF nodes than for task-level nodes. The

MCFs were not decomposed into task-level constructs because there was no data available to create a quantitative model at that level of detail. Similarly, no Crew Failure Modes (CFMs) were built into the model because there was insufficient information to support assigning specific CFMs from any taxonomy. As a result, this model includes only PIFs, MCFs, and HFEs as nodes.

5.4.1 Data Cleaning

The SACADA data available for this project was contained in spreadsheets, where one sheet recorded every MCF simulated by each crew along with its identifying information and the states of up to 29 PIFs. A second sheet provided more context to failed MCFs via an additional 23 PIFs. The first tasks of building the BN model from the SACADA data were interpreting and cleaning the data, and identifying a candidate case study scenario. For the purposes of this project, the data cleaning for model construction consisted of three major steps:

1. Replacing "Null" or "0" values in the "Cognitive Type" column with values assigned from keywords in the description of the MCF, via Table 5.1. This was done to ensure that cognitive phase-specific data was not simply thrown out and was assigned appropriately.
2. Removing extraneous, unnecessary, and/or completely null columns and rows from the data. Some columns (i.e., PIFs) were entirely null, in which case they were removed from the analysis. Information-only columns (tracking, e.g., year performed, scenario description, and experiment cycle) were removed for ease-of-use of the resulting dataset.
3. Removing special characters and simplifying the variable and state names for use in the Bayesian network software. Special characters (e.g., returns, punctuation) were common in the variable and state names of the original dataset. Removing them facilitated building the network nodes automatically from the data rather than having to do so manually.

Table 5.1: Keywords used to correct MCFs with unrecorded task cognitive types.

Keywords from SACADA TOE Column	Assigned Cognitive Type
Determine; Direct; Maintain; Monitor; Information; Refer; Detect	Information Gathering (1)
Enter; Evaluate; Relocate	Diagnosis or Decision Making (2)
Control; Act; Place; Perform; Remove	Action Taking (3)
Notification; Notify; Declare; Communicate; Notifies	Communication (4)

5.4.2 Data Mapping

Mapping the cleaned data to the HRA BN constructs consisted of creating a BN node for each SACADA element (PIFs and MCFs) that would be included in the model for the case study scenario. The MCF nodes were created by assigning variable names that indicated: 1) the cognitive type of the MCF, based on IDAC cognitive phases ([110–114]); and 2) the position within the scenario (i.e., assigning a number from 1 to *n* in order of performance in SACADA, where *n* is the number of MCFs in the scenario). For example, if the 6$^{\text{th}}$ MCF performed by a crew was Information Gathering (i.e., I-phase cognitive type), the corresponding MCF node in the model would be named "Information_Gathering_6." The states of the MCF were assigned as SAT (no failure) and UNSAT (failure) according to the designations in SACADA. The PIFs retained the name assigned in SACADA as the BN node name, and took the states assigned by SACADA.

The end states in SACADA tracked whether the crews succeeded (SAT) or failed (UNSAT) in the MCF, and also included grades on the success that indicated the crews performed better than expected (SAT+) or were deficient without strictly failing the MCF (SATΔ) [84]. For the purposes of this project, any deficiency was considered to be a failure, thus the UNSAT state in the model MCFs was informed both by the UNSAT and the SATΔ MCFs from SACADA. Incorporating SATΔ states as equivalent to UNSAT states ensures that any deficiency in performance is modeled appropriately, and helps to correct for the class imbalance (low-EPV nature) present in the SACADA data.

5.5 Method: PIF Selection

PIFs were assigned to specific cognitive phase MCFs, rather than assuming all PIFs were applicable to all MCFs. This assignment was based on the PIF name in SACADA, which indicated the applicable cognitive phase. This assignment method works for the purposes of this proof-of-concept. However, there are known issues with the SACADA dataset that limit the realism and applicability of the PIFs. Any PIF with "Overarching" in the name (e.g., Overarching Issues Time

Criticality) was assumed to be *a priori* applicable to all cognitive phases. However, not all PIFs from SACADA were included in the final BN model. This work included two main procedures for down-selecting PIFs for the final model, namely removing PIFs based on their 1) importance to determining the MCF state (UNSAT or SAT) and 2) Null or 0 value count.

5.5.1 Downselection via Feature Selection

In the course of building the model, it became apparent that some PIFs from SACADA were irrelevant or unimportant to their assigned cognitive phase MCF, in part because of the lack of causality in SACADA. This research focuses on causality, and therefore future models will represent the full, causal richness of PIFs carefully selected on the basis of scientific literature. However, since the current focus is on vetting the methodology rather than the final model, we have excluded some PIFs on the basis of their negligible mathematical impact. This was done because working with the smaller set of variables makes it easier to vet the method. As a result, it was determined that such "unimportant" PIFs should be removed from the modeling at this stage, since the goal was to invent the construction method rather than the final model. In final modeling, it is likely not necessary to perform feature selection at all.

To enable removing these PIFs, a *univariate feature selection* process was incorporated that would identify the *k* highest-scoring features (PIFs), i.e., those PIFs that were the most impactful for determining the state of the MCF. The process incorporated in this methodology was based on the univariate Pearson's chi-squared (χ^2) test to measure the degree of correlation between the PIF and the MCF. This therefore tested each PIF for its correlation/association to the MCF state. PIFs that scored highly in this test were then more important/impactful for determining the MCF state, and were retained in the model. The chi-squared test was effectively testing the veracity of the following hypotheses for each PIF:

- H_0: The PIF does not play a significant role in determining the state of the MCF.
- H_1: The PIF does play a role in determining the state of the MCF.

The chi-squared statistic is found by the summation across PIF states and MCF states of the difference between the expected frequency of the MCF being UNSAT when the PIF is in state i against the observed frequency of the MCF being UNSAT when the PIF is in state i. This test assumes that the frequency of the MCF being UNSAT will not be affected by the state of the PIF, i.e., the null hypothesis H_0 is that the PIF state is not a determining factor in MCF state. If the PIF state does not influence the MCF state strongly, the observed and expected frequencies will be similar, and the chi-square statistic (Equation 5.1) will be low. Conversely, if the PIF state is influential in determining the MCF state, the observed and expected frequencies of the MCF state will be significantly different, and the chi-square statistic will be high. In Equation 5.1, $E_{PIF=i}^{UNSAT}$ is the expected frequency of UNSAT MCF states when the PIF is in state i and $O_{PIF=i}^{UNSAT}$ is the observed (in the data) frequency of UNSAT MCF states when the PIF is in state i. Under the null hypothesis that the PIF does not play a role in determining the MCF state, the expected frequency in the Chi-squared test is equivalent to the marginal frequency of UNSAT MCF states.

$$\chi^2_{PIF} = \sum_i \left(\frac{\left(O_{PIF=i}^{UNSAT} - E_{PIF=i}^{UNSAT}\right)^2}{E_{PIF=i}^{UNSAT}} + \frac{\left(O_{PIF=i}^{SAT} - E_{PIF=i}^{SAT}\right)^2}{E_{PIF=i}^{SAT}} \right) \tag{5.1}$$

Using the chi-squared test therefore identified the k PIFs most important to determining the MCF state, and therefore the PIFs that should be retained in the model. The feature selection stage of PIF selection represents an important step in producing a robust model. This process, which must incorporate a methodology amenable to the data (e.g., the chi-squared based method for sparse data), focuses the model on the most important factors. Thus, feature selection works to optimize the tradeoff between model robustness and the resources required to build and quantify the model. For the purposes of this research, this procedure was implemented to automatically down-select to the 8 most important PIFs. This number was chosen to ensure model robustness (i.e., that there was sufficient explanatory power in the PIF set for each cognitive phase) while limiting the computational time and resources required to quantify the MCF CPTs. Including 8 ternary PIFs as parent nodes to a binary MCF would produce a CPT with 13,122 cells. It was decided that limiting the PIFs to those most impactful to the resulting MCF would produce a robust model with reasonable computational resources. In the case where a cognitive type did not have 8 PIFs to begin with, all PIFs were retained. However, this downselection process did not result in the final PIF set for each cognitive type. Rather, it was subsequently noticed that, despite their relevance in determining the MCF state, there were PIFs that were not amenable to modeling due to the presence of Null values in their CPTs.

It should be mentioned that other feature selection methodologies, beyond the chi-squared test, could be used. Particularly for categorical data (e.g., most HRA data), the Mutual Information between the PIF under study and its associated MCF could also be used to identify the highly-impactful PIFs. Additionally, feature selection methods based on Mutual Information can work both with categorical and numerical data, making Mutual Information a robust method regardless of the available data type [100]. This research, which used the available sparse data, incorporated the chi-square methodology because it uses observations in the data (and therefore accounts for sample size) rather than empirical distributions (as in Mutual Information) and is easily understandable. However, as datasets become less sparse for HRA, Mutual Information can be used for a robust feature selection process that does not require a specific data type. When Mutual Information was used to perform feature selection for this research, differences in the final PIF set were only found in the I and D cognitive phases, although none of the PIFs retained under chi-squared were excluded under Mutual Information.

5.5.2 Downselection via Null Value Count

As stated above, PIFs were assumed to be applicable only to specific cognitive phases, with the exception of the Overarching PIFs that were applied to all cognitive phases. However, some PIFs were populated with a large number of Null values which would have made a portion of the model essentially uncharacterized. Including these PIFs in the model would have increased complexity without a comparable increase in model utility. While there are methods available to correct for this sort of data imbalance, e.g., expert elicitation or Bayesian updating, devising such a scheme was beyond the scope of this research. Expert elicitation methods for augmenting data are discussed in

Section 2.5.1; a methodology for using Bayesian updating in HRA model construction is presented in [30]. The Null-heavy PIFs were instead identified and removed from the model.

Identifying the PIFs with a high percentage of Null values was relatively straightforward, and involved simply counting the number of Null entries as a fraction of the overall number of times the PIF was recorded and relevant. The overall length of the dataset (i.e., the number of total simulations recorded) was *not* used to determine which PIFs had a high percentage of Null values. This was because the PIF was expected to be Null for the performance of an MCF of a different cognitive type than was relevant for the PIF. For example, I-phase PIFs should have been Null for D-, A-, or C-phase MCFs, and therefore counting these Null values toward the relevance of the PIF to an I-phase MCF would be inappropriate. Accordingly, relevance was determined as the number of times a specific cognitive phase MCF was performed. For example, the fraction of Null entries for I-phase PIFs were only computed with respect to the total number of I-phase MCF performances. Equation 5.2 describes this process mathematically, where f_{Null}^{PIF} is the null-fraction of the PIF, $n_{PIF}(s_{PIF} = 0)$ is the number of times the PIF was recorded with a Null state during the performance of a relevant cognitive phase MCF, and n_{MCF} is the number of times the relevant cognitive phase MCF was recorded (i.e., the number of times the PIF ideally would be non-null):

$$f_{Null}^{PIF} = \frac{n_{PIF}(s_{PIF} = 0)}{n_{MCF}} \tag{5.2}$$

Table 5.2: Overview of the SACADA PIFs and the final set of BN PIFs. Cells marked with X indicate that the PIF was retained in the BN model for that cognitive phase.

PIF	States	PIF	States
I-phase		**D-phase**	
Alarm/Status Tile Detection Mode	Self-Revealing Procedure-Directed Check Procedure-Directed Monitoring Awareness/Inspection	Diagnosis or Response Planning (**X**)	Primarily Diagnosis Primarily Response Planning/ Decision Making
Alarm/Status Tile Status of Alarm Board	Dark Busy Overloaded	Diagnosis Basis	Procedure Skill Knowledge
Alarm/Status Tile Expectation of Alarm/ Indication of Change	Expected Not Expected Not Applicable	Diagnosis Familiarity	Standard Novel Anomaly
Meter/Light/Flag Detection Mode (**X**)	Procedure Directed Check Knowledge-Driven Monitoring Procedure-Directed Monitoring Awareness/Inspection	Diagnosis Outcome	Procedure-Based Activity Skill-Based Behavior Knowledge-Based Behavior
Meter/Light/Flag Individual Indicator (**X**)	Slight Change Distinct Change	Diagnosis Information Integration	Timing of Information Ambiguous Information Integration Required
Meter/Light/Flag Mimics/Displays/etc.	No Mimics Small Indications Similar Displays	Diagnosis Information Specificity	Specific Not Specific Not Applicable
Monitoring/Detection Detection Type	Alarm Status Tile Meter Indication Light Flag Computer Other	Diagnosis Information Quality	Missing Information Misleading Information Conflicting Information
A-phase		Decision Basis	Procedure Skill Knowledge
Type of Action (**X**)	Simple and Distinct Order Maintaining	Response Planning/ Decision Making Familiarity	Standard Adaptation Required Anomaly
Location (**X**)	Main or Aux. Control Board Back Control Panels	Response Planning/ Decision Making Uncertainty	Clear Uncertain Competing Priorities Conflicting Guidance
Guidance (**X**)	Procedure Skill of the Craft (Non-Faulted Hardware) STAR (Faulted Hardware)	Response Planning/ Decision Making Outcome	Procedure-Based Activity Skill-Based Behavior Knowledge-Based Behavior
		Overarching	
Recoverability (**X**)	Immediately Recoverable Recoverable with Sig. Effort Unrecoverable	Workload (**I, A, C**)	Normal Concurrent Demands Multiple Concurrent Demands
Additional Factors	Unintuitive Plant Response Unintuitive Controls Additional Mental Effort Reqd. Inadequate Feedback	Time Criticality (**I, D, A, C**)	Expansive Time Available Nominal Time Available Barely Adequate Time Avail.
C-phase			
Communication Driver (**X**)	Specifically Procedure Driven Not Specifically Driven	Extent of Communications Required	Nominal Communication Extensive Onsite Comm. Comm. within MCR
Communication Direction (**X**)	From (Simulator) Booth To (Simulator) Booth Public Address Announcement Other	Other Demands/ Factors	Non-Standard Noisy Background Coordination Communicator Unavailable Multiple Demands Memory Demands

Following this procedure, cognitive phase-specific PIFs were removed if the null fraction was over 0.5 ($f_{Null}^{PIF} > 0.5$), meaning the CPT would be quantified on 50% or less of the available data. This threshold was chosen because it was decided that, if more than 50% of the records were null, both the network validity and the ability to generate conditional PIF probabilities (to support dependency quantification) would be compromised. Overarching PIFs were removed using a similar procedure, but without considering the cognitive phase (i.e., the total fraction of Null values to performances was used). The output of this process was a set of PIFs with null fractions less than 0.5 for each cognitive phase, meaning the set of quantifiable PIFs relevant for each cognitive phase. This set of PIFs contained the final PIFs used for the BN. Table 5.2 shows the PIFs retained for each cognitive phase.

5.5.3 Final Node Set for Case Study Scenario

The nodes created by the previous section describe the general constructs of PIFs, MCFs, and HFEs usable in any HRA. For the purposes of this research, a single scenario was selected, an inter-system loss of coolant accident (ISLOCA) that was simulated by three crews and subsequently added to the SACADA database. Section 5.3 describes the scenario in more detail. The final nodes included for the case study are summarized and described below. While the nodes themselves may be scenario-specific, the same basic structure and node classes are usable across applications.

The case study scenario was composed of 13 distinct functions (MCFs), including at least one from each cognitive phase. Each function from SACADA was anonymized according to Section 5.4.2, but the PIF names from SACADA were retained.

The case study scenario included the "standard" nodes developed for the HRA constructs described previously (HFE, MCF, PIF), as well as some supporting structures that served to organize the model. The model includes "cognitive sub-phase" (CSP) nodes that serve as intermediaries between the MCFs related to a given cognitive phase (I, D, A, C) and the Objective node. The nodes can be broadly classified as found in Table 5.3. Beyond the nodes, the model was further organized into submodels, which are features of the GeNIe software that declutter the final model [76].

To improve model readability, each MCF and its associated PIFs were stored inside a submodel; this is a GeNIe structure that only serves to group related nodes and does not affect the inference process. Examples of the BN fragment for each cognitive phase (i.e., the nodes contained in each submodel) are shown in the following subsections.

Table 5.3: Overview of the node types available for use in an HRA BN, along with the number included in the case study model.

Node Class	States	Number in Case Study Model	Purpose
Objective	Binary: HFE; No HFE	1	Overall failure probability across the entire scenario.
CSP	Binary: HFE; No HFE	4	Failure probability attributable to a specific cognitive phase.
MCF	Binary: UNSAT; SAT	13	Failure probability of a single MCF.
PIF	Dependent: See Table 5.2	41	Causal factors influencing performance.
Filter	Binary: Null; Not Null	41	Re-parameterize PIF nodes to remove residual Null states from calculations. (Subsequently Removed).

5.6 Method: Cognitive Phase BN Submodel Creation

The following subsections provide the BN structures for each cognitive phase submodel, along with a brief explanation of the construction for each. Effect-modulating dependencies are indicated; more detailed discussion on assigning dependencies is provided in Section 5.7. Figures 5.1 – 5.4 are presented with evidence set on the Filter nodes. This reparameterizes the PIFs to remove the Null values and provide a clearer understanding of the probabilities associated with each PIF state (i.e., when not considering the Null state).

5.6.1 I-phase MCF BN Representation

The case study included one I-phase MCF during the scenario (*Information_Gathering_6*) along with three associated PIFs: Meter/Light/Flag Detection Mode, Meter/Light/Flag Individual Indicator, and Overarching Issues Time Criticality (see Table 5.2). The overarching PIF Overarching Issues Workload was also created as a parent to the I-phase MCF. Three filter nodes were added to the submodel, one for each PIF. The PIF and MCF nodes were created as discrete, probabilistic nodes in GeNIe; the MCF was built with states corresponding to SAT and UNSAT, and each PIF was built with the states assigned in SACADA. See Figure 5.1 for a visual representation of the I-phase submodel.

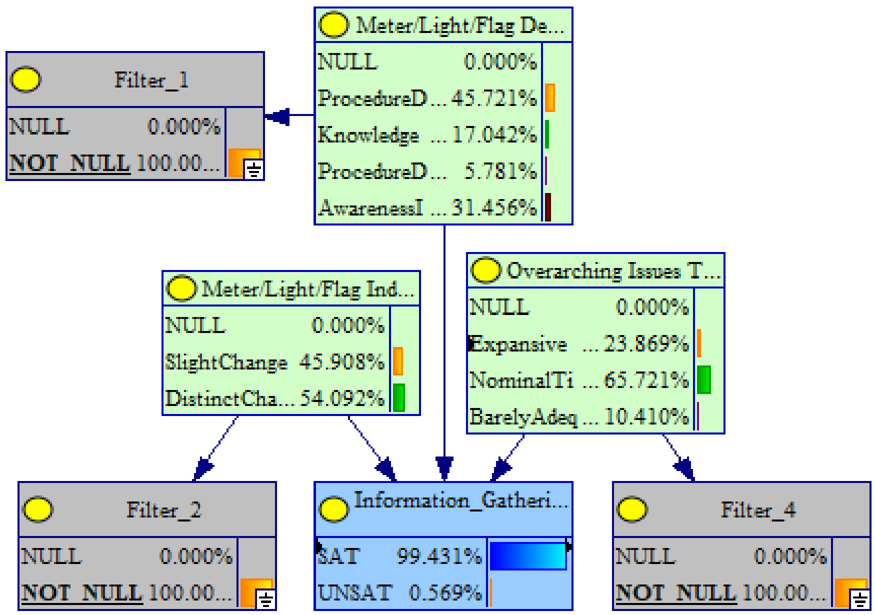


Figure 5.1: BN Fragment for I-phase MCF.

5.6.2 D-phase MCF BN Representation

The case study scenario included seven D-phase MCFs (MCFs 1, 3, 4, 7, 10, 11, 13), each with two associated PIFs: Diagnosis or Response Planning, and Overarching Issues Time Criticality (see Table 5.2). The overarching PIF Overarching Issues Workload was not included in the D-phase, because the PIF selection processes (Section 5.5) indicated that it was not important for D-phase MCFs. Two filter nodes were added to the submodel, one for each PIF. Note that there is an effect-modulating dependency between the D-phase PIFs. This was assigned on the basis that diagnosis and response planning tasks will take different amounts of time, which affects the ratio of time available to time required (i.e., Time Criticality). See Figure 5.1 for a visual representation of the D-phase submodel.

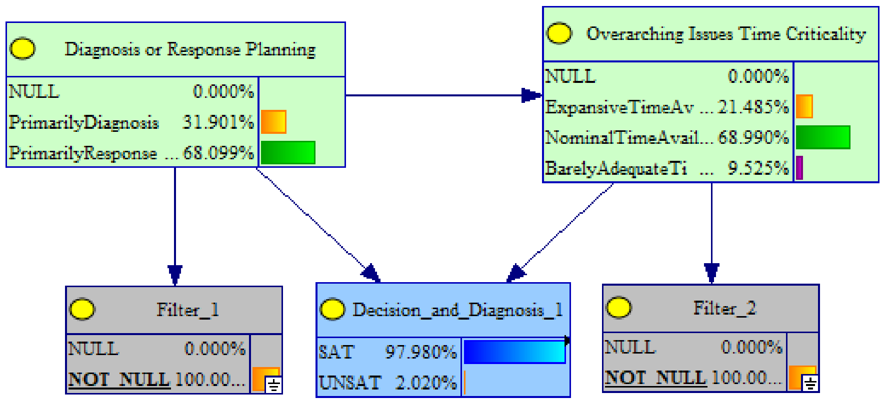


Figure 5.2: BN Fragment for D-phase MCFs.

5.6.3 A-phase MCF BN Representation

The case study scenario included four A-phase MCFs (MCFs 2, 8, 9, 12), each with five associated PIFs: Manipulation Type of Action, Manipulation Guidance, Manipulation Recoverability, Manipulation Location, and Overarching Issues Time Criticality (see Table 5.2). The overarching PIF Overarching Issues Workload was also assigned to the A-phase. Note that there are two effect-modulating dependencies in the A-phase submodel. Manipulation Type of Action and Time Criticality are assumed to interact because different types of actions will require different amounts of time, thus influencing the ratio of time available to time required (i.e., Time Criticality). Further, Manipulation Guidance was assumed to affect Manipulation Recoverability, because guidance for the recovery was assumed to be dependent on the guidance for the original task. Five filter nodes were added, one for each PIF. See Figure 5.3 for a visual representation of the A-phase submodel.

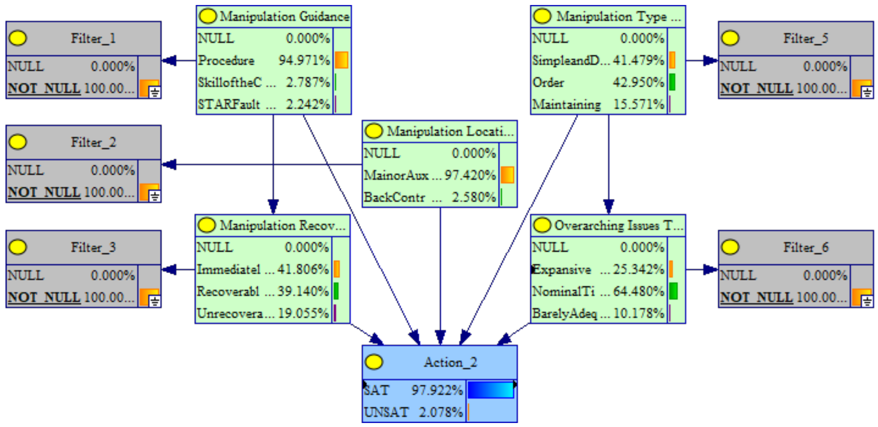


Figure 5.3: BN Fragment for A-phase MCFs.

5.6.4 C-phase MCF BN Representation

The case study scenario included one C-phase MCF (*Communication_5*) along with three associated PIFs: Communication Between Crew and Simulator Booth Direction, Communication Between Crew and Simulator Booth Driver, and Overarching Issues Time Criticality (see Table 5.2). The overarching PIF Overarching Issues Workload was also included as a PIF for the C-phase submodel. Note that there is one effect-modulating dependency in the C-phase MCF, between Communication Driver and Communication Direction. It was assumed that the presence of a specific directive for a communication would influence the direction of that communication. See Figure 5.4 for a visualization of the C-phase submodel.

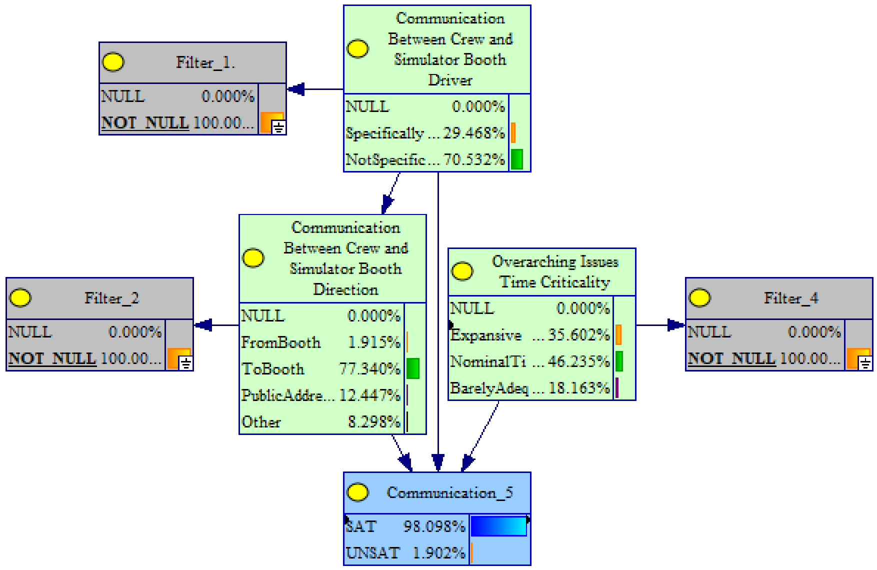


Figure 5.4: BN Fragment for C-phase MCF.

5.7 Method: Assigning Dependencies between HRA Constructs

The nodes created by the previous section, summarized in Figures 5.1 – 5.4 and Table 5.2, set up part of the structure for the network. However, creating the nodes of a Bayesian Network generated only half of the graph structure; without dependency relationships, there was nothing to drive the logic of the scenario. The next step after creating the BN nodes was to identify and encode dependency structures between the nodes.

5.7.1 Definitional Logic Dependencies

Dependency relationships with a definitional primary underlying logic, as defined in Section 4.4, include the definitional and construction idioms. Instances of definitional idioms were identified in the BN model built in this chapter. Instances of the construction idiom were not identified in the model because of a lack of available data for incorporating the idiom.

5.7.1.1 Definitional Dependencies

Definitional dependencies are defined in Section 4.3.2 and describe relationships wherein the parent node defines the child node, and therefore evidence of a parent node state precisely determines the child node state. These types of relationships are prevalent in *failure mode-function* variable pairs, where an observed failure mode *defines* the state (failed) of the associated function. In this work, definitional relationships were enforced between the MCF nodes and the "collector" node (cognitive sub-phase, CSP) that tracked the probability of MCF failure attributable to each cognitive phase. This follows the function-objective paradigm of definitional dependencies. If any MCF in a given cognitive phase was failed, the cognitive sub-phase node for that cognitive phase was defined as "HFE." Further definitional dependencies were assigned between each CSP node and the overarching Objective node, such that if any CSP node was observed to be HFE, the state of the Objective node was also HFE.

5.7.1.2 Construction Dependencies

Construction dependency idioms, as defined in Section 4.3.3, describe relationships wherein the parent node(s) collectively define the child node. However, in contrast to the definitional idiom, the relationship between parent(s) and child is *uncertain*. For discrete variables, this means that observing specific states of the parent(s) may not result in a certain observation of a single state of the child. Construction dependencies are therefore applicable when a construct is decomposed into multiple constituent pieces that together compose the construct, but when each piece has an uncertain relationship to the resultant construct. For example, "macro" PIFs such as Workload can be decomposed into, e.g., "Task Load" and "Non-Task Load."

There were no construction dependencies identified in the case study scenario under consideration. This is due to a lack of the data and/or information necessary to decompose extant SACADA structures into lower-level modeling constructs. For example, the PIF Overarching Workload provided in SACADA could have been decomposed into Task Load and Non-Task Load from the taxonomy in [42]. However, with no information on either Task Load or Non-Task Load for any of the MCFs, creating these new PIF nodes at a lower level of abstraction would have required making unsupported assumptions about the states. As will be discussed in Chapter 7, construction dependencies are useful for translating from low to high levels of abstraction (i.e., building a composite "Workload" node from constituent Task Load and Non-Task Load nodes), but the reverse requires additional information. As a result, construction dependencies were not included in this case study model.

5.7.2 Causal Logic Dependencies

Dependency relationships with a causal primary underlying logic, as defined in Section 4.4, include the causal, common context, situational, and effect-modulating idioms. Instances of all of these relationships were built in the BN model of the case study scenario.

5.7.2.1 Causal Dependencies

Causal dependencies, defined in Section 4.3.1, describe cause-effect relationships between two variables. In a BN, the parent node is the cause, while the child node is the effect. This is a probabilistic relationship, meaning that observing a given state of the parent (cause) affects a change in the probability of the child (effect) states, but does not precisely determine the child state.

In the model created for this project, causal idioms were constructed *a priori* between the PIFs and their associated MCF. It was assumed that causal relationships (at a minimum) were applicable between all of the PIFs in a cognitive phase and their associated MCF. For example, all PIFs relevant and important for I-phase MCFs were assigned as the causes of I-phase MCFs. Thus, the applicable PIFs for each MCF (see Table 5.2) were assigned as the parents of a causal dependency to their MCF. However, as other relationships were identified, some of the causal dependencies were replaced by effect-modulating dependencies. Figures 5.1 – 5.4 show the BN fragments for each cognitive phase, including the dependencies between the PIFs and the MCF. Note that in most cognitive phases, the strictly causal dependencies were replaced by effect-modulating dependencies, as the PIFs were determined to interact between themselves.

5.7.2.2 *Common Context Dependencies*

Common context dependencies, defined in Section 4.3.4, describe relationships between three or more variables. This dependency accounts for most of the "traditional" dependency factors considered in previous methods, e.g., THERP and SPAR-H. Common context dependency indicates that the children variables (i.e., MCFs) occur in a common environment represented by the parent variable(s), the PIF(s).

Assigning a common context dependency requires identifying PIF(s) that are persistent across a scenario, i.e., that represent factors that do not change in the relatively short timespan covered by the scenario. Such PIFs may be scenario-specific, as some scenarios (e.g., extended loss of offsite power (LOOP)) may take much longer than others to resolve. However, it can be generally posited that the PIFs that will drive common context dependency are those associated with aspects of the organization, (main control room) environment, and team cohesion; In most scenarios, such externally-defined factors will not change over the course of the relatively short time span. For the case study under consideration, which consisted of 13 MCFs performed inside the main control room, only one PIF was assumed to be static across the scenario. In this case, it was assumed that Overarching Issues Workload, which describes the number of tasks required to be performed, was a static attribute of the scenario. For this case study, Workload was found to be applicable for only I-, A-, and C-phase MCFs. As a result, the following MCFs were assigned as dependent through a common context of Overarching Issues Workload: *Action_2*, *Communication_5*, *Information_Gathering_6*, *Action_8*, *Action_9*, and *Action_12*. These dependencies are shown in Figure 5.6, where the Overarching Issues Workload PIF node is a parent to all the MCFs.

5.7.2.3 *Situational Dependencies*

Situational dependencies, defined in Section 4.3.5, describe relationships that occur *across time*. For discretely analyzed scenarios, MCFs are assumed to take place instantaneously and such that each MCF in sequence represents a distinct time step. Therefore, as in this research, "across time" should be taken to mean *across MCFs*, such that the relationship links two sequential MCFs together. Assigning a situational dependency requires understanding how subsequent PIF(s) can be affected by prior failures.

Perhaps the most salient example of situational dependency is a failure reducing the time available to perform the next MCF(s), and thereby increasing the probability of failure on those MCFs. For example, upon failing MCF A, a crew will do one of the following:

1. Recognize their failure on MCF A and thus commence recovery efforts prior to beginning MCF B,
2. Fail to recognize their failure on MCF A and thus incorrectly begin MCF B, or
3. Fail to recognize their failure on MCF A and thus perform incorrect subsequent MCFs (i.e., that are *not* MCF B).

In either case, the crew will have less time available to perform MCF B than would be provided by a successful completion of MCF A. Identifying, strategizing, and completing recovery actions take time, and thus necessarily reduce the time available for completing MCF B. On the other hand, if the crew fails to recognize MCF A failure, they will *incorrectly* start efforts on MCF B. However, because this means that their performance will not meet the system requirements, attempting MCF B after an unrecovered failure of MCF A is an error, and the time spent in this situation reduces the available time for the *correct* performance of MCF B.

Another salient example of situational dependency is a failure increasing the stress level on the HMT in subsequent MCFs. Failing an MCF A, for instance, may (upon recognition of the failure) increase the stress level of operators. This may occur because the operators are nervous about the scenario progression, upset with their error, or are now less familiar with the scenario. As a hormonal response, stress does not dissipate immediately, meaning that it will remain elevated for a time following the failure that caused the increase [37, 71]. Therefore, the initial failure increases the stress level of the operators for subsequent tasks, likely increasing the probability of failure.

Situational dependency may be particularly problematic when the initiating failure mode occurs in a D-phase MCF, e.g., “misdiagnosis” or “delayed diagnosis.” A-phase or C-phase MCFs may have more immediately-recognizable failures (e.g., miscommunication during three-way communication, incorrect action taken), which can be recovered in a short amount of time. However, D-phase MCF failures may be more difficult for the crew to recognize (and therefore more difficult to recover), and/or may require a more intense period of planning for how to recover, particularly if multiple subsequent tasks were performed based on the unidentified D-phase failure. As a result, failures in the D-phase may exert more drastic situational dependencies by greatly limiting the amount of time available for the correct performance of the following MCF(s).

In the case study network, the only situational dependency that was examined was the dependency from MCF A to Time Criticality to MCF B. Because Time Criticality was selected as a PIF for each cognitive phase type (Table 5.2), situational dependencies connect each MCF (i), $i > 1$, to its subsequent MCF ($i+1$). For the purposes of this research, it was assumed that Time Criticality was the only influential PIF in situational dependencies because the other PIFs represented attributes of the MCFs (e.g., Type of Action, Diagnosis or Response Planning, Detection Mode, etc.) that would not change based on previous failures. This point is elaborated in Section 5.9.5.

5.7.2.4 Effect-Modulating Dependencies

Effect-modulating dependency, defined in Section 4.3.6, describes a relationship between multiple PIFs that determine the non-linear probabilistic effects induced by combinations of PIF states. For instance, experiencing high stress concurrently with a limited amount of available time may result in a situation that is more error-prone than either PIF would suggest alone (e.g., experiencing *only* high stress or *only* limited available time). These are relationships that occur within a single time step (MCF) to exacerbate or mediate the effect of the environment on the probability of failure (similar to the error-forcing contexts in ATHEANA).

In the present case study, effect-modulating dependencies were identified in D-phase (Diagnosis or Response Planning, and Time Criticality), A-phase (Manipulation Guidance, and Manipulation Recoverability; Manipulation Type of Action, and Time Criticality), and C-phase (Communication Driver, and Communication Direction) MCFs.

5.8 Result: HRA BN Construction Methodology

The construction of BNs using the HRA dependency idioms as a guiding logical structure has multiple benefits, chiefly related to the traceability and causal focus of the HRA dependency idioms themselves. This section consolidates and generalizes much of the preceding chapter into a coherent methodology by which HRA data is transformed, through the use of the HRA dependency idioms, into a causal structure for HRA modeling.

Building the network begins by identifying the HRA constructs that will be included in the model, i.e., the MCFs, CFMs, and PIFs that are available in the data and pertinent to the scenario. These constructs are obtained from the HRA data, scenario artifacts (procedures, narrative, observations, etc.), PRA method guidance, and/or expert elicitation as available. If HRA data is used, the previous sections detail data cleaning and processing processes that are useful for translating the data into modeling constructs in the resulting BN. The use of expert-provided data (e.g., from an existing taxonomy of modeling constructs) and/or HRA methods may require mapping causal constructs, which will in turn require the development of appropriate methods for mapping.

The MCF nodes are constructed from the function- or task-level entities in the scenario, the exact level of abstraction of which may be determined by the data or PRA methodology used. Construction dependency can be used to construct MCFs from lower-level concepts (e.g., tasks and sub-tasks). An MCF should, as shown in Chapter 3, be an observable crew-system interaction. An MCF need not model a specific task; as in the case of the present research, MCFs may be used to model generalized IDAC cognitive phases. CFMs can be constructed from data and/or narrative associated with each MCF. If there is sufficient data and information to model the CFMs, they can be represented as individual nodes (with binary states relating to “Observed” and “Unobserved”), in which case each CFM becomes a definitional parent to its requisite MCF (which takes binary states relating to “Success” and “Failure”). In this case, the observation of any CFM will, through the definitional idiom, result in the “Failure” state of the MCF, and the the CFMs act as intermediaries between the PIFs and the MCFs. Alternatively, the CFMs may be directly represented as the states of an MCF, in which case the MCF will be multinary (with $n+1$ states: n CFMs and a Success state), and be directly connected to the PIFs, as in this research. If the information and data are insufficient to model CFMs, the definitional dependency between CFMs and MCFs allows for modeling the MCFs with binary states and directly influenced by the PIFs. The model developed in this dissertation uses only the PIFs and MCFs.

The PIF nodes are constructed from the non-task constructs in the HRA data and narratives by modeling orthogonal characteristics of the scenario related to the organization, environment, human-system interface, and crew personality. HRA datasets generally include PIFs explicitly ([84]), but PIF taxonomies ([42]) can also be used to identify and model PIFs from HRA narratives if required. The PIFs are modeled as multinary states (n states specific to each PIF).

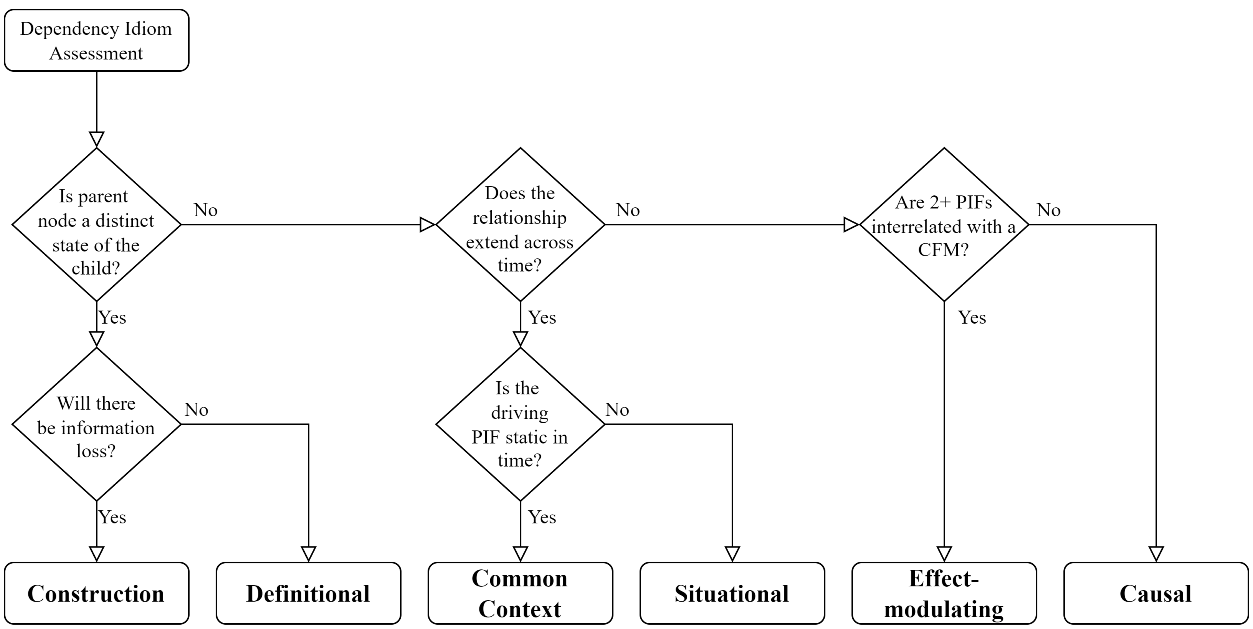


Figure 5.5: General methodology for assigning HRA dependency idioms.

The HRA dependency idioms are used to build the causal structure of the scenario once the nodes are created. The assessment of idioms to build the model structure generally follows the flowchart in Figure 5.5; more discussion on idiom construction is available in Section 5.10. Quantification is a separate process from idiom construction. Constructing the HRA dependency idioms entails building the arcs between the nodes and results in the full causal structure of the model.

This process, therefore, reveals the conditional probability tables (CPTs) that must be quantified, thus constraining the quantification process (Chapter 6).

Construction of the HRA dependency idioms should begin with the assessment of definitional dependencies (Section 4.3.2 that exist at least between CFMs and MCFs, as well as between the CSP (if included) and Objective nodes. The relationship between MCFs and CSP (or HFE) nodes will depend on the specific definition of HFE used in the model. For example, it may be desired that an HFE represents the failure of a single MCF, in which case the relationship will be definitional. Alternatively, if it is determined that the HFE will be defined by combinations of failed MCFs, construction dependency is used to connect the MCFs to the HFE node. In such a case, quantifying the construction dependency could use, e.g., k-out-of-N, OR, or Noisy-OR logic as deemed appropriate. Depending on the methodology and definition of the MCFs, definitional idioms may also exist between the MCFs and Objective node. In general, a definitional idiom is assigned between constructs wherein one (the parent) can be represented as a distinct state of the other (the child). Thus, definitional idioms occur from CFMs to their requisite MCF, because a CFM can serve as a distinct state of an MCF.

Causally-oriented dependencies should be enforced on the basis of viable causal mechanisms (available from literature), and/or be accompanied by documented justification. Causal idioms (Section 4.3.1) exist between each relevant PIF and its associated CFM (or MCF), where the PIF serves as a causal parent to the CFM or MCF. In this research, the PIF selection process (Section 5.5) produced a set of PIFs relevant to each MCF, and thus pre-determined the causal idioms in doing so.

Assessing common context dependencies requires first identifying PIFs that are both common to multiple CFMs or MCFs in the scenario and which are likely to (or can be assumed to) remain static across the scenario. Such PIFs are likely to include organizational factors and aspects related to the training, experience, and cohesion of the crew, which will not change in response to dynamic scenario changes. These PIFs represent the common context that connects multiple CFMs or MCFs across the scenario.

Situational dependencies should be enforced by inspecting each PIF in MCF i ($i > 1$) for a causal relationship to the CFMs (or MCF) in MCF *i-1*. That is, if a prior failure mode can influence the PIF, a situational dependency should be enforced by drawing a causal arc from the previous CFM to the PIF, which is already a causal parent of its CFM/MCF, thus enforcing a situational dependency. The PIFs involved in situational dependency are likely to include those PIFs that represent factors related to the task difficulty, local environment, and time available.

Effect-modulating dependencies reflect the interactions of multiple PIFs on the same CFM or MCF. Thus, assessing effect-modulating dependencies entails identifying those PIFs in the same submodel that are likely to interact, such as the interaction between task difficulty and time required. Because the PIFs are already encoded as causal parents to the CFM (or MCF), adding a causal arc between the interacting PIFs completes the effect-modulating dependency.

Construction dependencies are likely most applicable when building models at a higher level of abstraction than the data available. For instance, if the data provides low-level PIFs from the taxonomy in [42], but the model is using the higher-level PIFs, construction dependency can be used to build the higher-level constructs. Quantification of construction dependency is therefore construct-specific, and the subject of future research.

5.9 Result: HRA BN Built from HRA Dependency Idioms

The HRA dependency idioms facilitate a structured process by which data, expert elicitation, and literature can be fused to model the dependencies between different HRA constructs in a scenario. For example, expert elicitation can determine the dependency relationships present in the scenario, by investigating the HRA dependency idioms for their applicability to the scenario. This list is then verified and/or modified with available literature and further elicitation regarding the specific mechanisms that transmit each dependency, and pruned with preliminary data analysis to determine whether the dependency is supported by the data. Relationships that do not result in a significant probabilistic change in any variable may be discarded, simplifying the model with minimal impact to the final results. Finally, the dependencies are quantified with the empirical probabilities from the data.

The scenario of interest in the present case study included 13 MCFs, delineated between the IDAC cognitive phases per Table 5.4. Each TOE from the SACADA dataset was anonymized and modeled as only an MCF of the given cognitive type (see Figures 5.1 – 5.4) and assigned the relevant PIFs for that cognitive type (Table 5.2). Figure 5.6 depicts the full model representation in BN form.

Table 5.4: IDAC cognitive phase breakdown of case study MCFs.

IDAC Phase	**MCF Position in Model**	**Total Number in Model**
Information-gathering (I)	6	1
Diagnosis/Decision-making (D)	1; 3; 4; 7; 10; 11; 13	7
Action-taking (A)	2; 8; 9; 12	4
Communication (C)	5	1

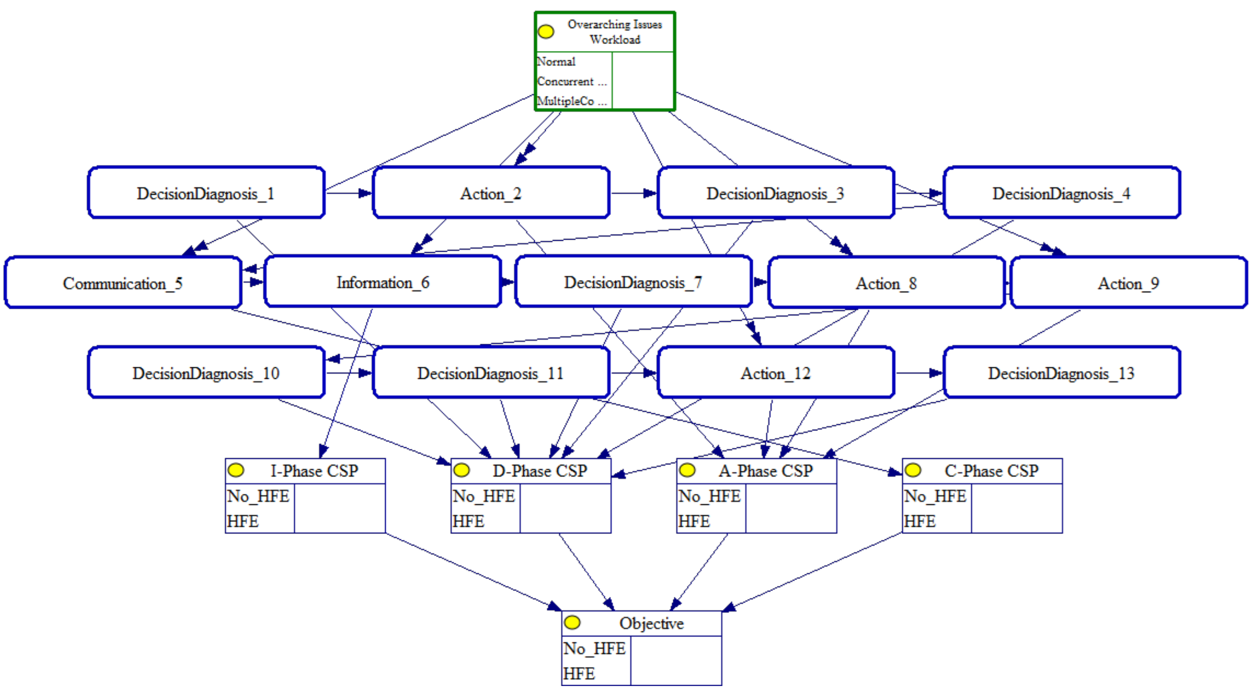


Figure 5.6: Full BN model of case study scenario.

The inductive construction of the network from HRA dependency idioms is an effective methodology for systematically considering and building dependency into the analysis prior to quantification. This enforces the use of *formative* rather than summative dependency and therefore minimizes the possibility that the analysis of dependency overrides other quantification of the model [15]. Inductive model construction follows a similar framework as most HRA methods, first identifying the tasks (MCFs) and factors (PIFs) relevant to each. However, whereas typical HRA methods might move into quantification of the probabilities as a next step (based on the PIFs and expert elicitation), the methodology presented above then searches for relationships between the constructs in the model. By examining which PIFs and MCFs (and CFMs) are expected to influence each other, using the HRA dependency idioms as a guide, the analysts are able to construct a fully-realized model of the scenario with dependency, before any quantification. This in turn will guide, and potentially simplify, the quantification process, as the required probabilities that must be gathered from data or expert elicitation are readily apparent. The idioms can therefore produce a causal model that is readily quantifiable.

5.9.1 Dependencies within I-phase MCFs

As Figure 5.1 shows, the I-phase MCF includes three PIFs (see Table 5.2) all exerting *causal dependencies* to the MCF node. The Time Criticality PIF and I-phase MCF are also dependent on the overarching PIF Workload. There is an effect-modulating dependency between Meter/Light/Flag Individual Indicator and Meter/Light/Flag Detection Mode, which was assigned on the basis that the magnitude of the indicator change would exacerbate or mediate the detection mode. For instance, a slight change in an indicator could complicate an already error-prone knowledge-based check or awareness-based inspection, while a distinct change in the indicator could make the task easier.

5.9.2 Dependencies within D-phase MCFs

As Figure 5.2 shows, the D-phase MCF includes two PIFs that are specific to the D-phase, and is also influenced by the overarching PIF Workload (see Table 5.2). The combination of the PIFs Diagnosis or Response Planning and Time Criticality with the MCF node form an *effect-modulating dependency*. This dependency was enforced based on the recognition that diagnosis functions and response planning functions may take different times, and that the time available may ease (extensive time) or complicate (limited time) both diagnosis and response planning functions.

5.9.3 Dependencies within A-phase MCFs

As Figure 5.3 shows, the A-phase MCF includes five A-phase specific PIFs, in addition to the overarching PIF Workload (see Table 5.2). There is a *causal dependency* between the PIF Manipulation Location and the MCF node. Additionally, there are two *effect-modulating* dependencies at work in this MCF. The PIFs Manipulation Guidance and Manipulation Recoverability form one effect-modulating dependency with the MCF node; the PIFs Manipulation Type of Action and Time Criticality form the second.

5.9.4 Dependencies within C-phase MCFs

As Figure 5.4 shows, the C-phase MCF includes three PIFs specific to the C-phase, as well as the overarching PIF Workload (see Table 5.2). There is a *causal dependency* between the PIF Time Criticality and the MCF node, as well as an *effect-modulating dependency* between the PIFs Communication Driver and Communication Direction with the MCF node.

5.9.5 Cross-MCF Dependencies

There are two types of dependency in the model that link separate MCFs to each other. There is a *common context dependency* driven by Overarching Issues Workload that connects the Time Criticality PIFs in I-, A-, and C-phase MCFs (see Table 5.4). This relationship was not included in D-phase MCFs because the PIF Overarching Issues Workload was not determined as relevant to D-phase MCFs. Additionally, there is a *situational dependency* between every MCF node to the Time Criticality PIF in the subsequent MCF; thus all MCFs are linked together via at least situational dependency.

The SACADA provided data includes some idiosyncrasies, including the assignment of PIFs to MCFs. As Table 5.2 shows, the PIFs can be broadly classified into factors inherent to the MCF, and factors that may change given the scenario. For example, a specific MCF (e.g., "Shut Valve A") has set guidance, occurs in a pre-defined location, and has a known recoverability. Previous failures will not typically affect these MCF-inherent PIFs. A previous failure will not change the amount of guidance for the MCF "Shut Valve A." Nor will a previous failure *necessarily* change the location in which the MCF will occur, although this may occur in some instances (e.g., if a previous failure disables a main control panel).

However, there are PIFs which previous failures will necessarily affect, such as Time Criticality. In SACADA, Time Criticality can be viewed as a ratio of the time required to perform a task (i.e., the time taken to perform), and the time available to perform the task. In any case, at least one of these aspects of Time Criticality is likely to be affected by previous failures. For instance, a previous failure will change the time available to perform the task, as part of the available time will be used to either recover the previous failure. A previous failure may also change the time required to perform the task, if environmental or mental conditions are altered to the detriment of performance.

For the case study presented herein, situational dependency was only assumed to act between MCFs through the PIF Time Criticality. This was necessary based on the PIFs included for each cognitive phase in the model, the majority of which were not influencable by previous MCFs. However, as the discussion above indicates, other situational dependencies are likely applicable in realistic scenarios, especially if additional PIFs are available for modeling purposes. With the anonymized scenario presented, the Time Criticality relationship was included because it does not rely on specific MCF details to ensure its inclusion.

As the preceding discussions indicate, the HRA dependency idioms created in Chapter 4 provide both the logic modeling structure for HRA dependency as well as a useful heuristic for creating causal BN models of a scenario. The heuristic for building the full BN from the HRA dependency idioms can also be used to trace existing dependency assessments. For instance, analysts could decompose existing full-scenario models according to the idioms to verify and validate the dependency relationships assigned between HRA constructs. This latter use-case could be strengthened with a comprehensive accounting of the causal mechanisms that underlie the HRA dependency idioms. However, identifying the specific causal mechanisms applicable between any two HRA constructs is beyond the scope of this work.

5.10 Analysis: Utility of HRA Dependency Idioms for Network Building

The HRA dependency idioms presented in Chapter 4 provide more than a useful graphical structure for HRA modeling. Perhaps more importantly, *the HRA dependency idioms structure the process of building the network* as well, by *providing a causal framework within which the scenario is analyzed*. Once the network nodes are determined (i.e., the MCFs, CFMs, and relevant PIFs are defined for the model), the causal structure is built by eliciting the idioms that are applicable between the nodes. Eliciting causal idioms within an MCF is a straightforward process; causal idioms are assumed *a priori* between the relevant PIFs and the associated MCF as part of the node selection process (i.e., irrelevant PIFs, which are not causal parents of the MCF, are not included in the model). Definitional idioms are similarly straightforward to elicit, as they are enforced (as a rule) between the MCFs to the requisite cognitive sub-phase node, and between the cognitive sub-phase nodes to the overall Objective node. Thus, the idiom elicitation process hinges on the assessment of common context, situational, and effect-modulating idioms between network nodes. However, as will be discussed, eliciting these idioms is made easier by following the graphical structures set out in Chapter 4, and assumptions made in the model structure mean that the elicitation process is relatively straightforward.

5.10.1 Common Context Idiom Assignment

Assigning common context idioms requires understanding the temporal nature of the PIFs present in the scenario. In general, there will be dynamic and static PIFs, referring to the behavior of the PIF states across the scenario. Understanding the PIFs that will not change during the scenario, because they are driven by scenario-external processes, will illuminate the common context idioms that are applicable in the scenario. Externally-driven PIFs are likely to include, for instance, PIFs related to organizational factors (e.g., management structure and work processes), procedures, and crew training/experience. Internal PIFs that vary across the scenario time could include PIFs related to the time available for a given MCF, the crew cohesion, and operator-internal PIFs such as stressors, awareness, and focus.

The externally-driven PIFs should then be investigated as drivers of common context dependencies, because they form a time-independent *context* in which multiple MCFs will be performed. For example, the procedures used in a scenario will not change across the scenario, and accordingly the MCFs will be dependent as a result of this context. Similarly, management structures and training/experience can impose common context dependencies due to the common administrative/process constraints and mental models that will govern the crew's performance in multiple MCFs.

Viewing the externally-driven PIFs as the drivers of common context dependencies in the network enforces the causal basis of these relationships. This is because, instead of viewing common factors (that may not be included as PIFs) as correlational aspects between two MCFs (as in previous HRA dependency methods), the common context dependency identifies the causal connections between the common PIF and each MCF. Identifying common context idioms requires finding which MCFs include the driving PIF as a causal factor, so pairwise investigation is sufficient to determine common context idioms. That is, it is sufficient to identify the cause-effect relationship from the common PIF to each MCF (or CFM, if available) to determine a common context idiom.

5.10.2 Situational Idiom Assignment

Assigning situational idioms similarly requires knowing the temporal nature of the PIFs in the scenario. Whereas common context idioms are enforced by the static, externally-driven PIFs, situational idioms work through the causal manipulation of dynamic, internally-driven PIFs that will change during the scenario. The state of an internally-driven PIF at a given point in time (e.g., during a given MCF performance) may be affected by previous failures, which is the essence of situational dependency. A previous failure may, for example, lower the time available for MCF performance, increase operator stress, and/or change how cues are presented.

Because the PIFs for a given MCF (i) are assumed *a priori* to impose causal dependencies on the MCF, eliciting a situational idiom requires identifying the PIFs in MCF i that are influenced by the failure or success of MCF (i-1). Thus, similar to common context idioms, situational idioms can be identified by finding the cause-effect relationship from MCF (i-1) to PIF(s) for MCF (i).

5.10.3 Effect-modulating Idiom Assignment

Effect-modulating idioms describe the nonlinear effects of PIFs acting in concurrently in time, meaning that identifying effect-modulating idioms requires identifying causal connections between PIFs within the same MCF. Identifying cause and effect in cross-MCF relationships (e.g., situational dependency) is eased by the unidirectionality of time, i.e., that the cause must precede the effect. However, this condition is not as visible when identifying effect-modulating relationships, which occur at the same time. As a result, eliciting effect-modulating idioms requires identifying and/or assuming a prime mover among the PIFs in a relationship, i.e. identifying which PIF in the effect-modulating relationship will be the parent of both directed arcs (see Figure 4.13).

This can be done by systematically examining the PIFs involved and understanding to which characteristics they correspond. For example, in the case study presented herein, there is an effect-modulating dependency between the A-phase PIFs *Type of Action* and *Time Criticality*. In this idiom, *Type of Action* is assumed to be the prime mover of the relationship. This is because the *Type of Action* for a given MCF is set externally to the scenario; the MCF "Shut Valve X" will always be a "Simple and Direct" task, just as "Maintain Departure from Nucleate Boiling" will always be a continuous, parameter-maintaining process. Thus, these are characteristics of the task itself, while *Time Criticality* is a characteristic of the task-scenario, meaning that it can be altered by developments internal to the scenario. The time criticality[3] in an MCF (i.e., whether the time

[3] Time Criticality in SACADA is a measure of both Time Available and Time Required

available is less than, equal to, or exceeds the time required for performance) has no bearing on the task type in the MCF – that is determined only by the MCF itself. However, the type of task (among other PIFs) may influence the time criticality. Thus, *Type of Action* is deemed to be the prime mover and *Time Criticality* the secondary PIF in this effect-modulating idiom.

5.11 Conclusions

This chapter created the methodology for constructing causal HRA BN models using the HRA dependency idioms. This is a significant advancement for HRA by improving the traceability of HRA model building and embedding dependency as the driving force of modeling (i.e., enforcing formative rather than summative dependency). This chapter shows how both HRA data and expert elicitation can be used in the construction of HRA BN models. Through the HRA dependency idioms, BNs constructed using this method encode causal understanding of how and why human error occurs in complex engineering systems. This chapter further demonstrates the methodology by building an HRA BN of an ISLOCA scenario from HRA data in the SACADA dataset. Using the HRA dependency idioms enforces traceability in the model construction.

Beyond the construction of an initial model, the BN submodels (Figures 5.1 – 5.4) also form reusable modeling structures that can inform the creation of subsequent models, as will be shown in Chapter 6. This is possible because the HRA dependency idioms model the reality of causal and definitional relationships, rather than application-specific connections. The idiom structures, because they model generalized relationships, do not change if specific functions are used instead of their respective cognitive phases. For instance, a specific Action function (e.g., "Start Safety Injection") used in place of a generalized A-phase MCF will still be subject to the effect-modulating dependencies from (Guidance and Recoverability) and (Type of Action and Time Criticality) and a causal dependency from Manipulation Location.

Using the HRA dependency idioms in this construction methodology provides a logical structure through which to analyze and assign dependencies, and the BN representation provides a visual indication of how dependencies manifest in the scenario. The logical structure for assigning dependencies means that the dependencies are elicited and verified on a pair-wise basis to iteratively and comprehensively build the full BN structure. These full BN structures are in turn the building blocks of HRA scenarios, which analysts can easily use in HRA practice. The logical process and BN structure let the analysts immediately identify whether the dependencies are possible and sensible, how the dependencies will affect the system, and ensure that the model structure is complete prior to quantification. Furthermore, basing the construction methodology on the HRA dependency idioms ensures that the process is traceable and scientifically valid, so that the final model is easily verifiable by external experts. The construction methodology developed in this chapter produces HRA BN models that support a deeper causal understanding of the system and scenario, provide a traceable representation of dependency, and allow for a manageable quantification process in Chapter 6.

Chapter 6: Quantifying HRA BNs with HRA Data

This chapter addresses the second portion of Research Objective 3, the development of the mathematics for quantifying dependency, including both the idioms and full-scope BN constructions. This chapter details the methodology for quantifying HRA BNs, such as the one built in Chapter 5, from HRA data while considering dependency at the forefront of the modeling process.

6.1 Introduction

HRA BN models created using the methodology developed in Chapter 5 are traceable and objective, but hitherto unquantified. Although such models provide robust qualitative understanding regarding the dependencies and causal factors affecting human error in complex engineering systems, quantifying the models (i.e., parameterizing the model CPTs) will enable the use of the models for quantitative HRA. Accordingly, this chapter creates a methodology for quantifying HRA BN models built from the HRA dependency idioms.

The emphasis of this chapter is on creating the methodology for quantifying HRA BNs with patterns of logic and data (which can include expert elicitation). Because this methodology is developed through a data-driven case study, this chapter also reviews the additional data cleaning that is required to support quantification. Further, this chapter creates the methodology for parameterizing model CPTs under both definitional and causal logic, when using HRA data (e.g., SACADA) to quantify the BN. Further, this chapter reviews processes for dealing with sparse and/or incomplete data, which is a common issue in HRA data.

While the quantification methodology itself is the principal result of this chapter, the quantitative results of the HRA dependency idioms and the CPTs of selected nodes are presented. These illustrate that dependency between HRA constructs is a real phenomenon with significant impact on the quantitative model, and it can be quantified using the methodology developed in this chapter. Finally, it is shown that end users can take models created from the methods in Chapter 5 and 6 and treat them like building blocks to expeditiously create necessarily models of different scenarios. The methodology developed in this chapter is a significant improvement over current HRA dependency quantification techniques. This methodology does not rely on the use of dependency levels and equations, as in previous methods for dependency assessment, but instead uses a model- and data-driven approach to dependency quantification. As a result, the dependency quantification methodology is traceable, objective and high-fidelity, and produces robust causal BN models to support quantitative HRA.

6.2 Approach: Data Cleaning for Quantification

As with the construction of the network in Chapter 5, the first tasks involved with quantifying the network (i.e., populating the conditional probability tables, CPTs) were to understand the data from SACADA and further modify it to suit the requirements of quantifying marginal and conditional probabilities. Much of the required data cleaning was performed prior to quantification, as discussed in Chapter 5. Building a *useful* quantitative model, however, required further data cleaning to expand the dataset, remove the filter nodes and quantify cross-MCF dependencies.

6.2.1 Data Rearranging

The chief aspect of data cleaning for network quantification is to transform the dataset into a two-dimensional array. In the initial data, each row did not necessarily correspond to a single crew performing a given MCF, but rather to the performance of that MCF by n crews ($n \geq 1$) under a nearly-identical PIF set (i.e., the PIF states recorded in one row corresponded to the performance of all n crews). There were some cases in which PIF differences were noted in the comments section, but generally single rows correspond to a nearly-identical PIF set for a single MCF performance by

n crews. In essence, the initial data could be viewed as a three dimensional array, with the MCFs and PIFs forming two axes and the number of performances forming the third axis. To facilitate quantifying the BN, this dataset was transformed such that each row constituted a single entry for each MCF as performed by each crew; see Figure 6.1 for a visual of this process. For example, if an MCF was performed n times, $n-1$ new copies of the MCF record were created and appended to the data array, so that the MCF was repeated in the final dataset n times.

Re-keying requires the appropriate ratio of "UNSAT" and "SAT" MCFs in each scenario to be encoded in the newly multiplied dataset. If an MCF is performed n times with an "UNSAT Ratio" of p (i.e., $100p\%$ of performances were failed), then (np) of the copies created in the previous step will be keyed as UNSAT, and $n(1-p)$ copies keyed as SAT.

MCF Name	Scenario Name	PIF 1 State		# Performed (n)	UNSAT Ratio (p)
MCF A1	Scenario A	PIF1_A1		3	1/3
MCF A2	Scenario A	PIF1_A2		2	1/2
MCF A3	Scenario A	PIF1_A3		2	0/2

Add (*n-1*) copies of each MCF row, where *n* is the number of times performed. *np* rows should be "UNSAT," the rest "SAT."

MCF Name	Scenario Name	PIF 1 State		UNSAT Ratio	UNSAT
MCF A1	Scenario A	PIF1_A1		0.3	1
MCF A1	Scenario A	PIF1_A1		0.3	0
MCF A1	Scenario A	PIF1_A1		0.3	0
MCF A2	Scenario A	PIF1_A2		0.5	1
MCF A2	Scenario A	PIF1_A2		0.5	0
MCF A3	Scenario A	PIF1_A3		0	0
MCF A3	Scenario A	PIF1_A3		0	0

Figure 6.1: The process of "multiplying" the initial SACADA dataset to produce a final dataset wherein each row contains only one performance.

6.3 Method: Quantification of HRA Dependencies

6.3.1 Quantification of Definitional Dependencies

Definitional dependencies rely on definitional logic and represent deterministic relationships between the parent and child nodes. Definitional dependencies were assigned between MCFs and their requisite cognitive sub-phase (CSP) node, as well as between each CSP to the final Objective node (see Figure 5.6). Construction dependencies were not included in the case study model. Quantifying the definitional dependencies was a straightforward task, because definitional dependency is a deterministic relationship. Thus, quantifying the definitional dependency involves only identifying the correspondence between parent and child node states. That is, identifying which parent node states define which child node states

Once the correspondence between parent and child node states is determined, the relationship can be quantified by setting the relevant child node state probability to unity. In the case study model, the failure of an MCF node (i.e., $MCF =$ "$UNSAT$") was assumed to result in the occurrence of an HFE in that cognitive phase (i.e., $CSP =$ "HFE"). Similarly, the observation of any CSP node as "HFE" was assumed to result in the final Objective node being "HFE." Mathematically, this means:

$$Pr(CSP_i = \text{“}HFE\text{”} | MCF_{i,j} = \text{“}UNSAT\text{”}) = 1, i \in \{I, D, A, C\}; j \in \{1, \ldots, n\} \tag{6.1}$$

$$Pr(Objective = \text{“}HFE\text{”} | CSP_i = \text{“}HFE\text{”}) = 1, i \in \{I, D, A, C\} \tag{6.2}$$

Similarly:

$$Pr(CSP = \text{“}HFE\text{”} | MCF = \text{“}SAT\text{”}) = 0 \tag{6.3}$$

Thus, quantifying definitional dependencies does not rely on HRA data to form the required conditional probabilities. However, it is significantly reliant on expert judgment to appropriately encode the definitional logic. For instance, an HFE may not be precisely defined as the failure of a single MCF. Rather, it may require the failure of multiple MCFs before an HFE is truly considered [88]. In such cases, the judgment of the analyst will be required to assign the appropriate states in the resulting CPT, although the quantification following this determination is not significantly more difficult than in the case previously discussed.

Construction dependencies were not able to be meaningfully incorporated into the case study model, due to the lack of sufficient data and information to do so. However, it is beneficial to theorize about the quantification of such dependencies nonetheless. Construction dependencies, as discussed in Section 4.3.3, are neither strictly causal nor strictly definitional relationships. Rather, they are a hybrid structure employing aspects of both to enforce an "uncertain definition" relationship. Construction dependency appears, for instance, when decomposing PIFs into lower-level constituents (or, alternatively, building low-level PIFs into macro-level constructs. As an example, consider the PIF Workload, which can be decomposed into Task Load (i.e., mission-critical actions or decisions) and Non-Task Load (i.e., required but non-mission critical actions and decisions, such as administration and communication) [42]. However, it is not clear precisely how the states of Task Load and Non-Task Load combine to form a cohesive unit in Workload.

Quantifying construction dependencies could be done using, for instance, k-out-of-N logic, noisy-OR logic or fuzzy set theory to quantify the conditional probabilities. A noisy-OR gate maps dichotomous parent nodes to a single child node, and includes a "leaky" aspect that allows for the observation of a given child node state even under contradictory evidence in the parent node states, thereby inserting a measure of uncertainty into the otherwise deterministic (definitional) relationship.

6.3.2 Quantification of Causal Logic Dependencies

Parameterizing these BNs requires using HRA information to parameterize the CPTs. Quantifying with data consists of first identifying the required empirical probabilities from data to produce a fully-quantified BN in which filter nodes (see Chapter 5) can be used to remove the residual Null states from the nodes. Subsequent sampling over the network is then used to re-quantify the CPTs without the filter nodes, which is discussed in Section 6.3.3.

This chapter presents the process for quantifying the built network fully from HRA data. The main approach used in this chapter is to directly compute the Maximum Likelihood Estimate (MLE) of the conditional probability directly from the data. We chose to use data rather than exploring data augmentation, expert elicitation, and/or parameter learning techniques; while these techniques could be used to counter some of the data issues related to sparsity and incompleteness, a full investigation into their utility is beyond the scope of the current work. The reader is referred to [30] for further discussion. In addition, Section 6.4.2 elaborates on the utility of the HRA dependency idioms for structuring an expert elicitation process, as well as the possibility of using the Expectation Maximization (EM) approach [115] for working with sparse and incomplete HRA data.

Now let us turn to the approach used to directly compute the Maximum Likelihood Estimate (MLE) of the conditional probability directly from the data. Quantifying any causal logic dependency entails determining the CPT associated with the child (effect) node, conditioned on each state of the parent node(s). The conditional probability of each state c of the child node C, conditioned on each state p of the parent node P can be found through the definition of conditional probability:

$$Pr(C=c|P=p) = \frac{Pr(C=c \cap P=p)}{Pr(P=p)}. \tag{6.4}$$

Equation 6.4 is the heart of the quantification process for all causal-logic dependencies. Translating to data-oriented notation, the joint probability in Equation 6.4 can be found by the definition of joint probability as the fraction of total data entries such that both C=c and P=p, where N is the number of data entries:

$$Pr(C=c \cap P=p) = \frac{N_{C=c,P=p}}{N_{total}}. \tag{6.5}$$

Similarly, the marginal probability in Equation 6.4 can be found as the fraction of total data entries such that P=p:

$$Pr(P = p) = \frac{N_{P=p}}{N_{total}}. \tag{6.6}$$

By combining Equations 6.5 and 6.6, it is shown that quantifying the conditional probability in Equation 6.4 from HRA data requires identifying the fraction of data entries such that P=p such that C=c is also true:

$$Pr(C = c|P = p) = \frac{N_{C=c,P=p}}{N_{total}} \cdot \frac{N_{total}}{N_{P=p}} = \frac{N_{C=c,P=p}}{N_{P=p}}. \tag{6.7}$$

The process developed in Equations 6.4 – 6.7 is central to the quantification of the other causally-oriented dependencies from data, particularly common context and situational dependencies. Because common context dependency is represented as a diverging BN structure, the CPT for each child node can be quantified independently of the other. Similarly, the chain structure of situational dependencies allows for the pair-wise quantification of the CPTs in the child and intermediate nodes. Constructing a CPT for a child node with multiple parents is done by expanding Equations 6.4 – 6.7 to incorporate multiple parents. For example, if a child node C has parents P_1 (with states p_1) and P_2 (with states p_2), the conditional probability of child node state $C = c$ is found (for each combination of parent states p_1 and p_2):

$$Pr(C = c|(P_1 = p_1; P_2 = p_2)) = \frac{Pr(C = C \cap P_1 = p_1 \cap P_2 = p_2)}{Pr(P_1 = p_1 \cap P_2 = p_2)}. \tag{6.8}$$

The joint probabilities in Equation 6.8 can be found in a similar manner to Equation 6.5:

$$Pr(C = c \cap P_1 = p_1 \cap P_2 = p_2) = \frac{N_{c,p_1,p_2}}{N_{total}}, \tag{6.9}$$

and:

$$Pr(P_1 = p_1 \cap P_2 = p_2) = \frac{N_{p_1,p_2}}{N_{total}}. \tag{6.10}$$

Thus, by substituting Equations 6.9 and 6.10 into Equation 6.8, the conditional probability of child node state c given parent node states p_1 and p_2 is obtained as:

$$Pr(C=c|(P_1=p_1,P_2=p_2)) = \frac{N_{c,p_1,p_2}}{N_{total}} \cdot \frac{N_{total}}{N_{p_1,p_2}} = \frac{N_{c,p_1,p_2}}{N_{p_1,p_2}} \tag{6.11}$$

For nodes with $n > 2$ parents, Equation 6.11 can be generalized to the chain rule of probability:

$$Pr(C=c|(P_1=p_1,...,P_n=p_n)) = \frac{N_{c,p_1,...,p_n}}{N_{p_1,...,p_n}} \tag{6.12}$$

The Equations 6.4 – 6.12 are relatively straightforward applications of the definition of conditional probability, and so this process is straightforward to implement in codes or software designed for data manipulation. Repeatedly grouping the data to find the joint probabilities is eased through the use of the Python package NumPy [116], which includes built-in functions to perform the data grouping and computation.

6.3.3 Filter Node Removal with Simulation

The network built in Chapter 5 included filter nodes on all of the PIFs in the network, due to the prevalence of "Null" or zero values in the SACADA dataset. The original purpose for the filter nodes was to allow evidence propagation without regard to the residual "Null" entries in the CPTs, which could insert erroneous results into the model. However, using the model with the included filter nodes would have required the analyst to set evidence on *every* filter node in the network, which can be a laborious task. As a result, it was necessary to prune the model and remove these nodes. This chapter developed the following process to create pseudo-data based on the built network:

1. Set evidence on filter nodes to "Not Null"
2. Sample over the network
3. Re-parameterize network with simulated data

These steps are useful in correcting for sparse and/or missing data, such as that in the currently-available SACADA dataset. As a result, these steps will not be generally required for BN modeling applications in HRA. Because this process samples over the pre-existing network, which was quantified with available real data, it will not significantly alter the quantitative results (i.e., will not alter the quantification of dependency). Similarly, this process will not alter the structure of the network, as the simulation requires the structure to generate meaningful results. For the purposes of this research, the sampling process described below was handled using GeNIe built-in functions [76], although this process can be performed manually. Sampling and re-parameterizing a network in this manner can be used when the network contains continuous nodes, which are parameterized with conditional distributions rather than CPTs, in which case it can be used to investigate the distribution tails with more detail.

6.3.3.1 Setting Evidence on Filter Nodes

The first step for data cleaning to remove the filter nodes was to set evidence on the filters themselves. This entailed setting each filter in the network to Non Null to ensure that the generated samples did not include any instances of any PIF being Null. The result was a data file of sample data points for each node in the network, wherein the data was entirely non-null. Therefore, this process corrected for missing and sparse data.

6.3.3.2 Sampling over the Network

With the evidence set on the filters, the next step is to iteratively sample from each node in the network. Because every node is included in each sample, this process can be computationally taxing for larger networks. Further, the computational effort will increase as the sparsity and/or incompleteness of the original data set increases. Because the process is meant to remove any null/zero values from the data, any sample that returns a null value will be removed. Thus, the percentage of samples that must be discarded to remove null values will increase as the initial data is more sparse or incomplete. For the case study in this work, roughly 18-20% of samples were

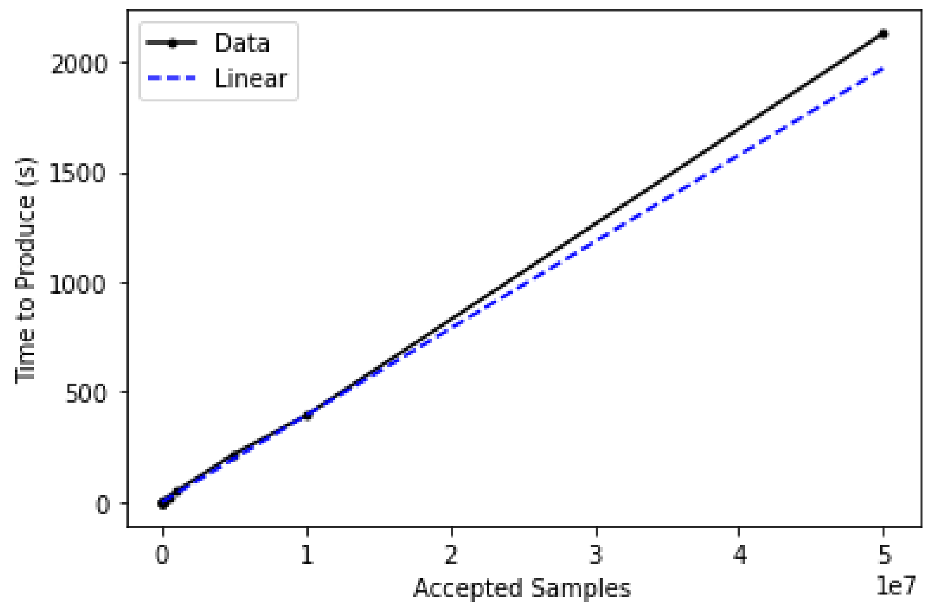


Figure 6.2: Sample generation time increases linearly with the number of samples demanded.

discarded, i.e., 5,000,000 samples were generated in total to produce 1,000,000 accepted samples. Varying the desired number of accepted samples for the network in this project from 100 to 50,000,000 reveals that an average of 18.94% of samples were discarded during the sampling process. The discard fraction does not substantively change as more accepted samples are demanded; the sample standard deviation of the discard fraction is 0.0014 as shown in Table 6.1. The time required to generate the samples increases roughly linearly, as shown by Figure 6.2.

Table 6.1: Overview of discard fraction and time required for generating up to 10,000,000 samples.

Accepted Samples	**Rejected Samples**	**Discard Fraction**	**Time Required (s)**
100	428	0.1894	0.001
1,000	4,389	0.1856	0.06
10,000	42,322	0.1911	0.53
100,000	425,183	0.1904	5.23
1,000,000	4,287,133	0.1891	49.27
10,000,000	42,745,863	0.1896	393.66
	Average	0.1894	
	Sample Standard Deviation	0.0014	

6.3.3.3 *Re-parameterizing the Network*

The data produced by sampling over the network in the previous step produces an array of size (s,n) where s is the number of accepted samples and n is the number of nodes in the network, meaning the array is likely to be extremely large. Thus, the computational time that will be required to extract, analyze, and re-parameterize the network with the data should be considered when determining the number of samples to generate. In the present work, the resulting data files for 10,000,000 samples over the 100-node network (including filters) ranged from 2 to 13 GB in size. Such file sizes posed a limitation for using GUI-based platforms (e.g., MS Excel or the native GeNIe viewer) and increased the time required to analyze the data using Python.

The generated data file facilitates quantifying the network directly via the process described in Section 6.3 or via algorithms such as Expectation Maximization (EM) [115, 117]. This research used both direct quantification and the EM algorithm (through a built-in GeNIe function) applied to the sampled dataset, which returned nearly identical results. The removal of filter nodes does not change the causal structure between the HRA modeling constructs that was created using the HRA dependency idioms in Chapter 5.

6.4 Result: Methodology for Dependency Quantification

The previous sections detail the processes for parameterizing the CPTs with HRA data, by using the HRA dependency idioms as the guide to understanding the probabilities that must be obtained from the data. For example, the effect-modulating dependency in the D-phase submodels (Figure 5.2) indicates that marginal probabilities must be found for the PIF Diagnosis or Response Planning, while conditional probabilities are required for both the PIF Overarching Issues Time Criticality and the MCF. As a result, the HRA dependency idioms clearly show which parent nodes must be accounted for in the conditional probability parameterization, e.g., that the state probabilities of the PIF Overarching Issues Time Criticality must be conditioned on each state of its parent, the PIF Diagnosis or Response Planning.

Finding the requisite probabilities can be done using the data driven strategy presented below (and explained in the previous sections) or via a process of expert elicitation which will similarly be structured by the HRA dependency idioms.

6.4.1 Data-driven Strategy

HRA data is typically composed of discrete variables. As a result, quantifying the CPTs for the variables is a matter of finding the conditional probabilities, as shown above. Parameterizing the CPTs for a BN in turn quantifies the effects of the dependency relationships that are encoded in the network. Figure 6.3 shows the general process for parameterizing a CPT. Quantifying a CPT for a q-state node with m parents, each having s_i states ($i \in \{1,...,m\}$, requires quantifying $q\prod_i^m s_i$ cells in the CPT. As a result, the data must be grouped iteratively such that each combination of possible parent states is represented; this process is performed automatically by the Python grouping functions [116].

The process for quantifying CPTs in BN nodes is a fairly straightforward application of basic laws of probability, as previously explained. With data available, there is no elicitation of dependency levels or calculation of dependency modifier that is required. Instead, structuring dependency relationships using the idioms allows for the direct computation of the dependency based on the data. Further, the idiom dependency structure shows the set of immediate parent nodes for any dependent (child) node. Quantifying dependent nodes relies *only* on the direct parents, due to the chain rule of probability for BNs (Equation 6.13), which shows that the joint probability of a group of nodes can be determined as the product of each node conditioned on its respective parents. This also indicates that the CPT of any given node only relies on its parents, as any parent node will be similarly conditioned on its parents.

$$Pr(X_1, X_2, ..., X_N) = \prod_i Pr(X_i | pa(X_i)) \tag{6.13}$$

The data-driven quantification process developed above does not suffer from the time and effort

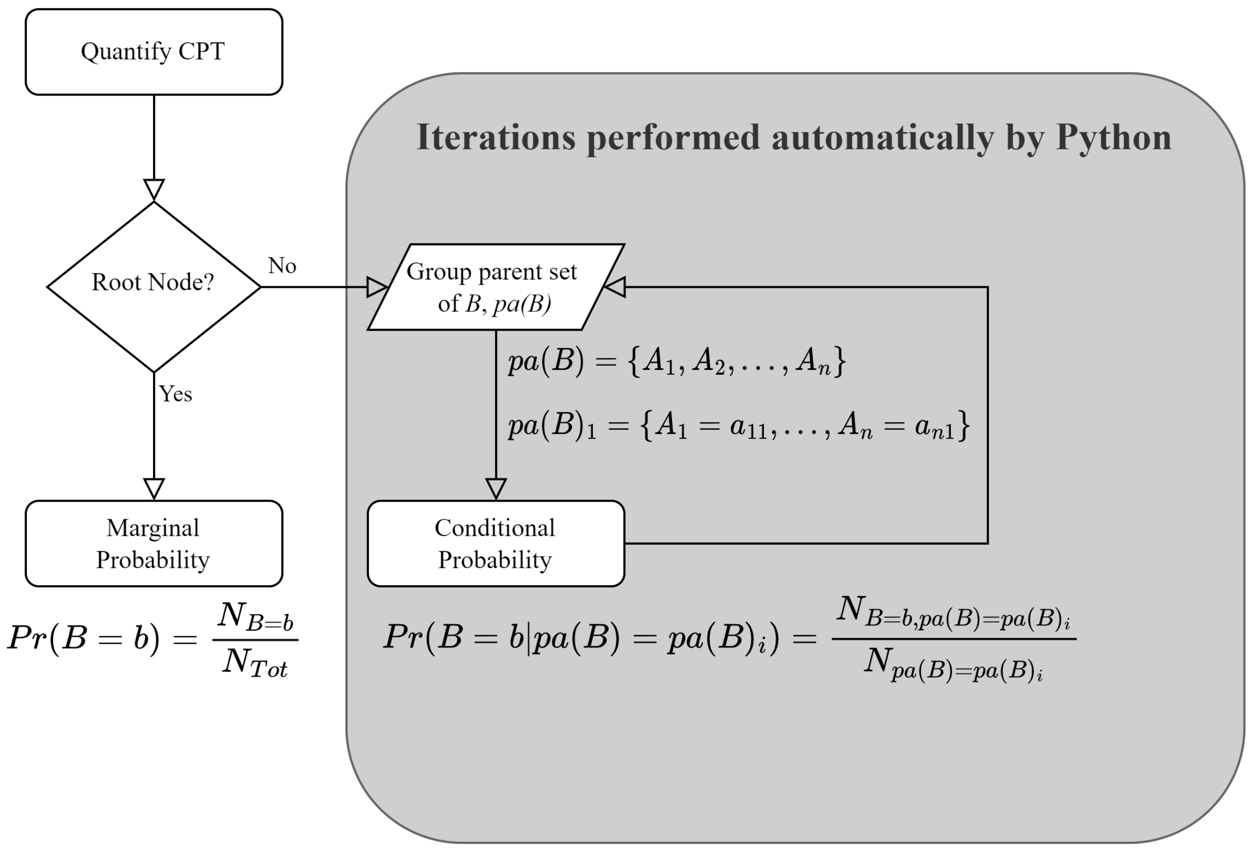


Figure 6.3: Flowchart for Quantifying CPTs in HRA BN.

constraints imposed by, for example, expert elicitation process, especially when automated through available software (e.g., Python). The idioms discretize the dependency structure, which allows for efficient parameterization via the chain rule of probability (Equations 6.12, 6.13). Therefore, each node CPT can be parameterized by grouping the set of immediate parents and iteratively finding the probability of the child node state conditioned on all combinations of the parent nodes. This is a fully model- and data-driven process that handles sparse and incomplete data using the laws of probability and without necessitating expert elicitation. Accordingly this methodology is *traceable* and *objective* through the idiom structure and laws of probability, and *high-fidelity* through its basis in HRA data rather than the use of dependency levels and equations.

6.4.2 Alternative Quantification Strategies

For highly-incomplete data, such as is common in data sparse contexts like HRA, the network can be improved via re-quantification techniques like iterative sampling (Section 6.3.3) to correct null/missing values. This method only retains samples where no null values are recorded, and so the computational time required will scale with the sparsity of the data. If data is incomplete or extremely sparse, parameter learning algorithms such as Expectation Maximization (EM), can be used to parameterize the CPTs [117]. A thorough discussion of the EM algorithm is beyond the scope of this research, although a brief overview is provided below. For more information on the use of EM for BN quantification, readers are referred to [115, 117]. This research determined that a robust iterative sampling technique achieves results similar to the EM technique, although this may not be true in cases of extreme sparsity.

Directly quantifying the model from data, using the definitions of marginal and conditional probability, is only one available method for quantification. Although it is an objective and straightforward process that will result in a fully data-driven model, it may not always be possible to implement. In the severe absence of data, or in the presence of different modeling requirements, there are alternative methods that can be employed in substitution to, or augmentation of, a direct data-driven quantification scheme [118]. These alternative quantification methods include *expert* *elicitation*, wherein persons with expertise in the requisite area provide values for the probability in question, either directly or indirectly. Expectation maximization is an algorithm that can compute the maximum likelihood estimates of BN parameters from incomplete data, such as that available in HRA and other applications [115, 117].

6.4.2.1 Expert Elicitation

Expert elicitation methods allow analysts to assign conditional probabilities on the basis of expert knowledge regarding the relationship in question. While it is possible to elicit precise probabilities, this is not recommended as it can induce biases into the quantification [119]. As a result, there are structured methods available for eliciting "fuzzy" probabilities, mapping linguistic assessments to probability ranges, and combining expert judgments; some of these methods are the focus of ongoing research, as discussed in Section 2.5.1. Expert elicitation is a valid method of parameterizing BN CPTs, and is particularly useful in the absence of data. However, structured processes for elicitation (e.g., the Delphi Method) and/or aggregating elicitations (via, e.g., Analytic Hierarchy Process (AHP) or Analytic Network Process (ANP)) should be used to correct for individual biases in each expert and arrive at a more objective value for the model. Further, a diverse group of experts should be consulted to remove systemic biases that may be present in specific expert groups (i.e., one should consult experts in the design, operation, and maintenance of a system).

In the case that expert elicitation is used to parameterize the BN model, using the idioms allows for a further structuring of the elicitation process that, when combined with the methods discussed above, can produce a robust, traceable, and objective quantitative model. In the idiom-based expert elicitation approach, the HRA dependency idioms indicate the required probabilities, and the conditional probabilities associated with each idiom can be individually elicited. For example, parameterizing the CPTs in a D-phase submodel (Figure 5.2) with expert elicitation would entail separately eliciting the marginal probability of each state of Diagnosis or Response Planning, as well as the probability of each state of Time Criticality conditioned on each state of Diagnosis or

Response Planning. Thus the HRA dependency idioms provide the structure for a traceable and robust expert elicitation process and clearly indicate the probabilities that must be elicited.

Expert elicitation has been widely incorporated in previous HRA methods, and current research is focused on developing more advanced techniques to improve the traceability and objectivity of the elicitation process, as discussed in Section 2.5.1. These methods, which can produce robust values that minimize personal and systemic biases in the estimates, should be incorporated in an expert elicitation process for quantifying HRA BNs.

6.4.2.2 Expectation Maxmization

Expectation maximization is an algorithm that iteratively computes the maximum likelihood estimates for BN parameters, which are typically the parameters of conditional distributions that define continuous BN nodes. However, it is possible to implement this process with discrete data, in which case the conditional probabilities are computed directly. The expectation maximization algorithm consists of two steps (expectation and maximization) in an iterative process that ultimately converges at the maximum likelihood estimate of the desired parameter [115]. The EM algorithm begins with specification of an initial value of the parameter of interest ϕ (i.e., the required conditional probability). Then, the *E-step* and *M-step* are performed iteratively until convergence. Each *E-step* (iteration i) estimates a sufficient statistic t value (t^i) by finding the expected value of the statistic given the current value of parameter, ϕ^i. Subsequently, the *M-step* determines a new value of ϕ, ϕ^{i+1}, that maximizes the likelihood of observing t^i. That is, the *M-step* determines the maximum likelihood estimator of ϕ, which becomes ϕ^{i+1} and the next iteration begins. This process eventually converges to the maximum likelihood estimator for ϕ, and works with incomplete data. For a thorough explanation of the expectation maximization algorithm, see [115, 117].

6.5 Result: Quantified Dependency for HRA

The BN for the case study scenario was built using the dependency idiom structure and quantified directly from HRA data available in SACADA [84]. *The numerical results shown in this* *section are not meant to replace HEP values that are currently used in HRA, nor are these results meant to formulate new dependency multiplier values for use in, e.g., Equation 2.1. The numerical results of this work are meant to show that a coherent, useful BN including formative dependency can be built from the HRA dependency idioms and HRA data, and that dependencies have a demonstrable effect on the numerical results.* In this case, "coherent" means that the baseline (i.e., marginal) human error probabilities are both reasonable and in general agreement with existing HRA assumptions and knowledge. The general agreement of the baseline results with current HRA practices further emphasizes the utility of the network for examining the probabilistic changes engendered by the HRA dependency idioms.

6.5.1 HRA Dependency Idiom Quantification Results

The network built for the case study modeled MCFs as general instances of IDAC cognitive phases (i.e., as I-, D-, A-, or C-phase MCFs instead of specific MCFs, see Figure 5.6) [110–114]. The model was quantified with the full SACADA dataset, meaning that the probabilities are not specific to the scenario at hand. This allows for comparison of this model to assumed and known baseline HEP values, such as those used in THERP and SPAR-H, to show the acceptability of the model values. Table 6.2 compares the baseline HEP values for each IDAC cognitive phase; Table 6.3 shows the comparative effects of different HRA dependency idioms on the resultant HEP.

As Table 6.2 shows, the baseline values obtained from this model align reasonably well with previous methods and assumptions in HRA. Further more, these values align with the marginal HEPs computed directly from SACADA, as the ratio $\frac{(UNSAT+SAT\Delta)}{Total}$ and those values published from SACADA [120]. This means that the construction of the network did not impose any errors and that the SACADA data is generally aligned with HRA assumptions and principles, and therefore that a comparison of the model conditional probabilities (due to different HRA dependency idioms) to the model baseline HEPs is not erroneous.

Table 6.3 shows the changes in the HEP for each cognitive phase engendered by specific causal, common context, and effect-modulating dependencies. In general, the results are sensible in that

Table 6.2: Comparison of baseline HEPs from this model to values used in THERP and SPAR-H.

	Baseline HEP Value		
Cognitive Phase	**THERP [16]**	**SPAR-H [31]**	**This model**
Information Gathering	< 0.1*	—	0.0060
Diagnosis & Decision-making	0.0001 – 0.1**	0.0100	0.0200
Action Taking	< 0.05***	0.0010	0.0202
Communication Between Crew	0.001 – 0.2****	—	0.0191

* THERP Tables 20-9 to 20-11
** THERP Table 20-1, assuming >10 minutes available
*** THERP Table 20-12, outlier value of 0.5 discounted
**** THERP Table 20-8, column (b) assuming detailed orders

PIF states anticipated to increase the HEP do typically increase the HEP as expected. Having barely adequate time to perform the function increases the HEP in all cognitive phases, particularly in A-phase MCFs, which is sensible considering the likely variance in time required among action tasks (compared to, e.g., information-gathering or diagnosis tasks). Having multiple concurrent tasks similarly increases the HEP in I-, A-, and C-phase MCFs, consistent with the idea that multiple tasks present distractions. Finally, the non-linear effects of multiple negative PIF states are apparent as PIFs interact to form a context that is more conducive to human error than either PIF state would indicate separately. Note that D-phase MCFs are not dependent on the workload; this may be because I-, A-, and C-phase MCFs still have elements of a physical "action" required (e.g., having to go and search for information in the I-phase) that can be affected by having to do other tasks, while D-phase MCFs are distinctly cognitive functions. Alternatively, high-level decision making may be performed by the Shift Technical Supervisor (or equivalent) and therefore less affected by workload.

Table 6.4 depicts the changes in HEP produced by situational dependency from the previous MCF ("Driving MCF" in Table 6.4), through Time Criticality to the MCF in question ("Subsequent MCF" in Table 6.4). This idiom is meant to capture the situation wherein a previous failure reduces

Table 6.3: Comparison of model baseline HEPs to dependency-borne conditional HEPs.

Cognitive Phase	**Model Baseline HEP**	**Causal (Time Crit. = *Barely Adequate*)**	**Common Context (Workload = *Multiple Concurrent*)**	**Effect-modulating**
I	0.0060	0.0069	0.0113	0.0384*
D	0.0200	0.0261	N/A	0.0446**
A	0.0202	0.1334	0.0314	0.1102***
C	0.0190	0.0271	0.0264	0.0277****

* Time Criticality = *Barely Adequate*; Workload = *Multiple Demands*
** Time Criticality = *Barely Adequate*; Diagnosis Task
*** Time Criticality = *Barely Adequate*; Type of Action = *Order*
**** Communication Driver = *Not Specifically Directed*; Direction = *To Booth*

Table 6.4: Review of Situational Dependency Quantitative Results

Driving MCF	**Subsequent MCF**	**Baseline Probability of Subsequent MCF**	**Conditional Probability of Subsequent MCF (previous failure)**	**Magnitude of Change (%)**
I	**D**	0.0202	0.0205	1.5354
D	**D**	0.0202	0.0205	1.4851
D	**A**	0.0208	0.0207	-0.4808
A	**D**	0.0202	0.0205	1.4851
A	**A**	0.0208	0.0209	0.4808
C	**I**	0.0057	0.0081	42.1053
			Mean Change (%)	**7.7685**
			Median Change (%)	**1.4851**

the time available to perform the subsequent MCF. The theory behind this idiom is that, if MCF *i* is failed, the crew may use time allotted to the performance of MCF *i+1* to diagnose, plan, and recover from their previous failure.

It should be noted that many SACADA scenarios are training scenarios, meaning that the scenario is designed to be more difficult (i.e., more error-inducing) than "typical" contexts encountered during normal operations, which may account for some of the variation between the model values and the representative values from previous HRA methods. Further, SACADA published values do not treat SATΔ records as failures, which may be the cause of some of the variation between model baseline results and SACADA values found in [120]. However, the general agreement in marginal probabilities to values used in mature HRA methods points to the fact that the quantified network is sufficiently reliable to draw *some* preliminary quantitative conclusions regarding dependency.

6.5.2 Quantitative Dependency Results for IDAC Cognitive Phases

Tables 6.5 – 6.8 summarize the HEP changes enforced by the causal dependency idiom from each PIF parent node. As with Tables 6.3 and 6.4, these results highlight the utility of the causal idiom for producing reasonable (i.e., aligned with expectations) quantitative results, and are meant to show the efficacy of modeling dependency with the idioms rather than provide values to replace existing dependency modifiers.

Table 6.5 shows the change in HEP (from marginal/baseline to conditional) produced by a causal dependency from each PIF state to the associated MCF. A representative baseline HEP for I-phase MCFs is 0.0060. Note that, generally, the PIF states expected to increase the HEP do increase the HEP (albeit with varying magnitude); for instance, Barely Adequate Time Available increases the HEP, as does Knowledge-driven Monitoring. Note that "Procedure Directed Monitoring" drastically increases the HEP, which may be indicative of issues with the availability and/or quality of the procedures. This supports the recognition that poor-quality and/or unavailable procedures are a large driver of human error [1]. Therefore, the I-phase MCF fragments behave generally as expected when quantified with SACADA data, and the included dependencies are impactful for changing the HEP.

Table 6.6 shows the effects of causal dependency from each PIF to D-phase MCFs. A representative baseline HEP for D-phase MCFs is 0.0201. For D-phase MCFs, the PIF states expected to increase the HEP (e.g., Primarily Diagnosis, Barely Adequate Time Available) increase the HEP with respect to the baseline value. Therefore, the D-phase fragments behave generally as expected when quantified, and the dependencies included in the D-phase are valuable for determining the HEP.

Table 6.7 shows the effects of causal dependency on the A-phase HEP from each associated PIF.

Table 6.5: Quantitative results for causal dependency on I-phase MCFs. A representative baseline HEP for I-phase MCFs is 0.0060.

I-phase PIF State	Conditional HEP	I-phase PIF State	Conditional HEP
Meter/Light/Flag Individual Indicator		**Meter/Light/Flag Detection Mode**	
Slight Change	$\downarrow$ 0.0056	*Procedure Directed Check*	$\downarrow$ 0.0049
Distinct Change	$\uparrow$ 0.0064	*Knowledge Driven Monitoring*	$\uparrow$ 0.0173
Overarching Issues Time Criticality		*Procedure Directed Monitoring*	$\uparrow$ 0.2294
Expansive Time Available	$\downarrow$ 0.0005	*Awareness Inspection*	$\downarrow$ 0.0015
Nominal Time Available	$\uparrow$ 0.0070		
Barely Adequate Time	$\uparrow$ 0.0069		

Table 6.6: Quantitative results for causal dependency on D-phase MCFs. A representative baseline HEP for D-phase MCFs is 0.0201.

D-phase PIF State	Conditional HEP	D-phase PIF State	Conditional HEP
Diagnosis or Response Planning		**Overarching Issues Time Criticality**	
Primarily Diagnosis	$\uparrow$ 0.0282	*Expansive Time Available*	$\downarrow$ 0.0163
Primarily Response Planning or Decision Making	$\downarrow$ 0.0162	*Nominal Time Available*	$\uparrow$ 0.0211
		Barely Adequate Time	$\uparrow$ 0.0261

Note that the presence of procedural guidance increases the HEP, both compared to the marginal value and the values conditioned on the presence of other types of guidance. One might expect procedure-based tasks to have a lower HEP, because the crew can identify, perform, and then verify each distinct step with the procedure. However, Table 6.7 reveals that the opposite is true in this model. Because the marginal HEPs are relatively aligned to both SACADA nominal values and typical HRA assumptions, and other PIF states adjust the HEP as expected, this may be indicative of unexpected confounding variables. For instance, low-quality or unavailable procedures are a concern for human error [1], and there is no variable that represents the quality or applicability of guidance. What may be occurring in this instance is that procedure-based tasks are more likely to have poor-quality guidance than, for instance, skill-of-the-craft actions, meaning that the presence of a procedure may actually impart a higher probability of error. Interestingly, having expansive available time also increases the HEP, which is unexpected. This result may be a product of more difficult functions generally being associated with longer performance times (and, therefore, more time allotted), but being more prone to failure by virtue of their difficulty.

Table 6.7: Quantitative results for causal dependency on A-phase MCFs. A representative baseline HEP for A-phase MCFs is 0.0204.

A-phase PIF State	Conditional HEP	C-phase PIF State	Conditional HEP
Manipulation Location		**Manipulation Recoverability**	
Main or Auxiliary Control Boards	$\uparrow$ 0.0208	*Immediately Recoverable*	$\uparrow$ 0.0212
Back Control Panels	$\downarrow$ 0.0081	*Recoverable with Significant Effort*	$\downarrow$ 0.0146
Manipulation Type of Action		*Unrecoverable*	$\uparrow$ 0.0293
Simple and Direct	$\downarrow$ 0.0149	**Overarching Issues Time Criticality**	
Order	$\uparrow$ 0.0261	*Expansive Time Available*	$\uparrow$ 0.0831
Maintaining	$\downarrow$ 0.0195	*Nominal Time Available*	$\downarrow$ 0.0161
Manipulation Guidance		*Barely Adequate Time*	$\uparrow$ 0.1350
Procedure	$\uparrow$ 0.0212		
Skill of the Craft Non-faulted Hardware	$\downarrow$ 0.0034		
STAR Faulted Hardware	$\downarrow$ 0.0108		

Table 6.8 shows the probabilistic impacts of causal dependencies on C-phase HEPs. A representative baseline HEP for C-phase MCFs is 0.0191. The results in this phase also generally align with expectations: PIF states anticipated to produce higher HEP values do tend to produce higher HEPs. For instance, Barely Adequate Time Available increases the HEP, as does the absence of specific procedure directions for communication. The C-phase BN fragments therefore behave generally as expected with respect to the effects of PIF states on the HEP. However, additional caution should be exercised when examining C-phase results, because of the general lack of data regarding C-phase TOEs in SACADA (C-phase TOEs are significantly underrepresented in the already sparse dataset).

6.5.3 Selected Conditional Probability Tables

The results discussed in the previous subsections are the descriptive statistics that show the general effects of dependency on the conditional probabilities in the network. However, descriptive statistics are just that: descriptive, and not thorough. The Tables 6.3 – 6.8 provide a good summary of the dependencies for each cognitive phase, however additional insight for both dependency and data improvements can be gleaned by investigating the conditional probability tables. Tables 6.9

Table 6.8: Quantitative results for causal dependency on C-phase MCFs. A representative baseline HEP for C-phase MCFs is 0.0191.

C-phase PIF State	Conditional HEP
Communication Between Crew and Simulator Driver	
Specifically Procedure Directed	↓ 0.0064
Not Specifically Directed	↑ 0.0245
Overarching Issues Time Criticality	
Expansive Time Available	↑ 0.0805
Nominal Time Available	↓ 0.0183
Barely Adequate Time Available	↑ 0.0271
Communication Between Crew and Simulator Direction	
From Booth	↑ 0.1474
To Booth	↑ 0.0203
Public Address	↑ 0.0193
Other	↓ 0.0062

– 6.11 show the complete CPTs for PIFs and the MCF in a D-phase submodel (Decision and Diagnosis_4, specifically). Table 6.12 provides the partial CPT for the D-phase cognitive sub-phase (CSP) node, which is definitionally dependent on the D-phase MCFs. Appendix B provides representative CPTs for nodes in the other cognitive phases, as well as the CPTs for overarching nodes in the model. Note that in the following tables, the conditional probabilities have been truncated from the original values in GeNIe.

Table 6.9 shows the CPT for the PIF Diagnosis or Response Planning, which describes whether the MCF in question related more to the diagnosis of an identified anomaly or the planning of a response. Because this node is a root node (i.e., no parents), the probabilities are the marginal probability of each state of the PIF. This CPT indicates that, in the SACADA dataset, 68 % of D-phase MCFs relate to response planning or decision making, while 32% relate to purely diagnosis tasks.

Table 6.9: CPT for the PIF Diagnosis or Response Planning in the D-phase submodels.

PIF State	Probability
Primarily Diagnosis	0.3194
Primarily Response Planning or Decision Making	0.6806

Table 6.10 displays the CPT for the PIF Overarching Issues Time Criticality, which describes the ratio of required performance time to the available time. This PIF is dependent on both Diagnosis or Response Planning (in an effect-modulating dependency) and the previous MCF (in a situational dependency). Note that, generally, diagnosis tasks are associated with a higher probability of expansive available time than decision making or response planning tasks. Additionally, the CPT shows the changes in expected time available as the result of previous MCF outcome (i.e., a previous MCF failure is associated with lower probability of expansive available time).

Table 6.10: CPT for the PIF Overarching Issues Time Criticality in the D-phase submodels.

PIF State	**Conditional Probability**			
Diagnosis or Response Planning	*Primarily Diagnosis*		*Primarily Response Planning or Decision Making*	
Previous MCF	*SAT*	*UNSAT*	*SAT*	*UNSAT*
Expansive Time Available	0.2683	0.2296	0.1882	0.1862
Nominal Time Available	0.6451	0.6699	0.7116	0.7206
Barely Adequate Time Available	0.0865	0.1005	0.1002	0.0931

Table 6.11 shows the CPT for the MCF node in the D-phase submodels, which describes the probability of a success (SAT) or failure (UNSAT), conditioned on the two D-phase PIFs (Diagnosis or Response Planning; Overarching Issues Time Criticality). Note the significantly higher probability of UNSAT (failure) associated with diagnosis tasks performed under barely adequate time, a product of the effect-modulating dependency.

Table 6.11: CPT for the MCF in the D-phase submodels.

PIF State	**Conditional Probability**					
Diagnosis or Response Planning	*Primarily Diagnosis*			*Primarily Response Planning or Decision Making*		
Overarching Issues Time Criticality	*Expansive*	*Nominal*	*Barely Adequate*	*Expansive*	*Nominal*	*Barely Adequate*
SAT	0.9886	0.9665	0.9588	0.9818	0.9848	0.9817
UNSAT	0.0114	0.0335	0.0412	0.0182	0.0152	0.0183

Table 6.12 shows the partial CPT for the D-phase cognitive sub-phase (CSP) node. This node

is definitionally dependent on all D-phase MCFs, and as a result only takes the state "No HFE" if all D-phase MCFs are SAT (i.e., there is no D-phase HFE). If any D-phase MCF is failed, the D-phase CSP node takes the state HFE. Thus, when no evidence is present, the D-phase CSP node indicates the total probability of experiencing a D-phase HFE in the scenario.

Table 6.12: Partial CPT for the D-phase CSP node.

CSP State	**Conditional Probability**								
D-phase MCF (1)	*SAT*								$\rightarrow$
D-phase MCF (3)	*SAT*								$\rightarrow$
D-phase MCF (4)	*SAT*								$\rightarrow$
D-phase MCF (7)	*SAT*								$\rightarrow$
D-phase MCF (10)	*SAT*				*UNSAT*				$\rightarrow$
D-phase MCF (11)	*SAT*		*UNSAT*		*SAT*		*UNSAT*		$\rightarrow$
D-phase MCF (13)	*SAT*	*UNSAT*	*SAT*	*UNSAT*	*SAT*	*UNSAT*	*SAT*	*UNSAT*	$\rightarrow$
No HFE	1	0	0	0	0	0	0	0	$\rightarrow$
HFE	0	1	1	1	1	1	1	1	$\rightarrow$

6.6 Discussion: Quantitative Insights for HRA Dependency

This chapter created a traceable, objective, and high-fidelity methodology for quantifying dependency in HRA. This is a step forward in understanding and quantifying dependency, which has typically relied on arbitrary equations instead of traceable, causal dependencies robustly quantified. As a result, this method enables rigorous modeling of dependency in HRA based on a more scientific understanding of how and why human error occurs, as well as an understanding of how the causal constructs in HRA can influence each other. This methodology leverages the HRA dependency idioms and laws of probability to traceably and objectively glean the necessary probabilities from data (including expert knowledge). The methodology was implemented on a case study to demonstrate how it can be used to create first-of-kind quantified, formative dependency models for HRA.

The quantification methodology is the key result of this chapter, but the numerical results of the case study model are illustrative. Notably, the numerical results of the modeling indicate that dependency in HRA does indeed impact the overall results, and can be quantified with the HRA dependency idioms. This chapter has shown that dependency is a real and impactful phenomenon that occurs between multiple HRA variables/constructs that can be quantified with these newly introduced methods. This chapter has developed a traceable and objective methodology for quantifying dependency in HRA models. Due to the sparsity and incompleteness of the SACADA data, future work is necessary to apply this methodology to create more robust quantitative models (e.g., [121, 122]) and gather more HRA data to support modeling.

Examining the model suggests that such models could be used to gather new insights into the nature of HRA dependency[1]. Of particular interest are the magnitudes of the effect-modulating dependencies (Table 6.3), which show that interacting PIFs can significantly change the resulting HEP compared to either PIF alone. For example, if both *Barely Adequate* Time Available and the Action is an *Order*, the conditional HEP is 0.1102 (11.02%). However, if only *Barely Adequate* Time Available exists, the conditional HEP is 0.0477 (4.77%); *Order* Actions alone have a conditional HEP of 0.0272 (2.72%). The quantitative results from causal and common context dependencies are similarly clear, each resulting in substantial changes to the HEPs in multiple cognitive phases.

Quantifying situational dependency is a more difficult proposition than quantifying the non-temporal dependencies, as it relies on sequential data to compute the probability of a variable state based on the previous task parity (i.e., success or failure). With sparse data, there may be insufficient instances of previous MCF failures to gain meaningful insight into how previous failures change the probability of subsequent PIFs, as was true in this research. Still, the numerical results from quantifying situational dependency (see Table 6.4) indicate that situational dependency does occur between subsequent MCFs. Further, the strongest (i.e., most impactful to the subsequent HEP) situational dependencies appear to be from D-phase functions to A-phase functions. This is

[1]Caution: the exact value of the probabilistic changes in the case study model are not meant to replace existing dependency modifiers; they are provided to show that the method is a viable way to build full models in future work.

consistent with the idea that, if the preceding diagnosis or decision is incorrect, time allotted for the subsequent action will be spent trying to identify and resolve the error. Thus, a failed decision/diagnosis should increase the HEP on a subsequent action, which is precisely what the data indicates. Additionally, there is also evidence for strong situational dependencies between A-phase functions, which is similarly consistent with the recognition that the available time for the second action will be used to recover from the failed first action.

Table 6.4 indicates that the situational dependencies between other cognitive phases result in lower, but still observable, changes in the subsequent HEP. The variability in results should serve as the impetus for future research into the causal mechanisms that result in situational dependency between functions. This will benefit from additional data, although it may require some restructuring of the data to better support modeling efforts. Situational dependency is not explicitly accounted for in current HRA methodologies, yet as this research has shown, it may have real and significant effects on the overall HEP, because of relationships that can be substantiated with real mechanisms. Situational dependency

The dependency quantification for each cognitive phase revealed results that generally adhered to expectations regarding the effects of various PIF states. However, the C-phase results are complicated by additional factors compared to the other cognitive phases. In the original SACADA dataset, only about 48% of TOEs were assigned a cognitive type (the non-keyed TOEs were assigned a cognitive type based on keywords from Table 5.1), and of these only 7.84% were C-phase functions. As a result, C-phase data is not only sparse but heavily underrepresented in the overall data, and so additional data collection should be performed to better understand C-phase tasks.

The models built using the comprehensive methodologies created in this dissertation are thus fully-quantifiable HRA models with formative dependency built into the model structure. There are useful insights that can be gleaned from the case study model. For example, there are interacting PIF states that will exacerbate or mediate each other to form more or less error-forcing conditions, as predicted with effect-modulating dependency, defined in Section 4.3.6. Further, the causal, situational, common context, and effect-modulating dependency relationships can support both positive and negative dependency (i.e., different states will increase or decrease the HEP), meaning that the idioms can support positive and negative dependencies between multiple HRA constructs, including cross-MCF relationships and the effects of PIFs. Finally, evidence on a single construct can be propagated across the entire network, meaning that unexpected causal pathways can be elucidated through investigating the network itself, and predictions of HMT performance throughout the scenario can be made. The methodologies presented herein therefore created a full HRA model that can capture all relationships of interest both broadly in HRA and specifically for the definition of dependency produced in Chapter 3.

The case study BN model therefore represents the entirety of the ISLOCA scenario. Further more, the submodels representing each cognitive phase can be re-used as building blocks to form the model for a new scenario. This was tested by building a new model for a fire-HRA scenario also contained in the SACADA dataset [84]. This is discussed briefly in the next section, which shows the ability to construct new models from the building blocks built from the HRA dependency idioms.

6.7 Discussion: Extending the Idioms to a Second Scenario

One of the advantages of the HRA dependency idiom structures is their reuse ability, particularly when implemented with generalized constructs (e.g., cognitive-phase MCFs instead of specific functions) as building blocks of HRA. To understand the viability of reusing the HRA dependency idiom structures as building blocks to produce additional models, the following steps were performed to construct a distinct BN for a separate scenario, using the same idiom structures for each cognitive phase (i.e., Figures 5.1 – 5.4):

1. Identify an applicable new scenario
2. Recreate the idiom structure of each cognitive phase
3. Reconnect cross-MCF dependencies
4. Verify network structure
5. Copy CPTs from quantified idiom structures to applicable CPTs in new network

6.7.1 Identify Applicable Scenario

The first step to extending the idioms to additional scenarios was to identify an applicable scenario to which the existing idiom-based BN fragments are amenable. For the purposes of this research, this was simply identifying another scenario in the SACADA dataset, since the included scenarios were recorded generally at the same level of abstraction with the same PIFs. However, building a model from scratch (for, e.g., a prognostic application in a novel scenario) may require defining the modeling constructs at the same level of abstraction as the pre-built idioms and BN submodels support.

The scenario identified for this example extension was a Fire Response scenario recorded in SACADA, which featured the following sequence of cognitive phase MCFs:

1. Decision & Diagnosis
2. Decision & Diagnosis
3. Action Taking
4. Decision & Diagnosis
5. Action Taking
6. Action Taking
7. Decision & Diagnosis
8. Action Taking
9. Communication
10. Action Taking
11. Action Taking
12. Action Taking
13. Information Gathering

Note that this scenario is more action-oriented, while the original ISLOCA scenario used for the case study included many more cognitive-focused (i.e., D-phase) MCFs, as shown in Section 5.3. However, this difference in scenarios does not limit the applicability of the idioms and pre-built BN fragments to modeling for the Fire Response scenario. One of the only limiting factors in this application was the presence of particular cognitive patterns that exist in the Fire Response model, but not in the ISLOCA model. For example, the MCF cognitive patterns $A \rightarrow C$ and $A \rightarrow I$ are both present in the Fire HRA network, as $MCF\ 8 \rightarrow MCF\ 9$ and $MCF\ 12 \rightarrow MCF\ 13$, but are not present in the ISLOCA scenario. This may present a challenge to quantification for situational dependency, but not to model construction.

6.7.2 Recreate the Idiom Structure for Each Cognitive Phase

Constructing the model for the new scenario from the idiom structures built for existing models entailed recreating the idiom structures and BN fragments from the existing network and sequencing them to match the new scenario. In GeNIe, this was done by copying the idioms and submodels from the network built for the ISLOCA to a new model, and then renaming and/or reorganizing the new model constructs as appropriate. For example, in creating the Fire Response model from the idioms and BN submodels constructed for the ISLOCA model, all that was required was copying the requisite BN fragments (i.e., Figures 5.1 – 5.4) from the ISLOCA model (Figure 5.6) into the new model for the Fire Response scenario (Figure 6.4).

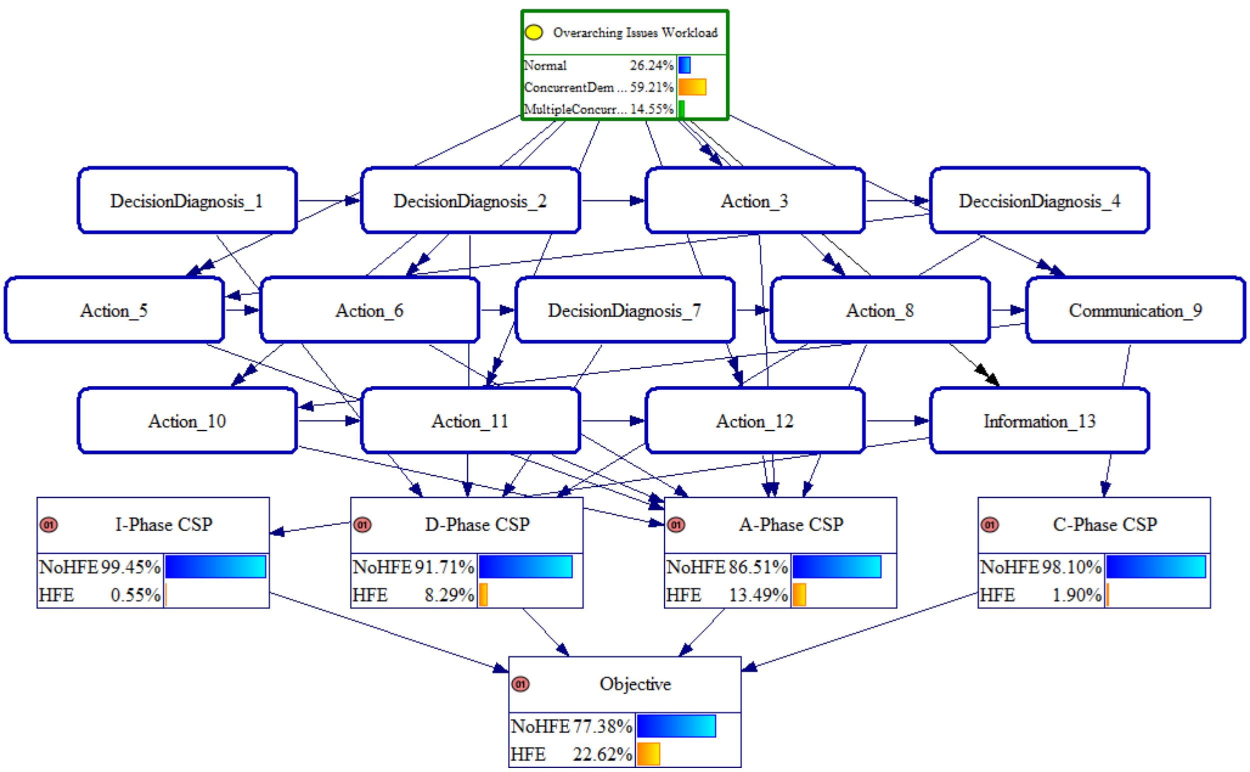


Figure 6.4: BN for the Fire Response scenario created with the idiom structures and BN building blocks from the ISLOCA model.

Because the idioms represent the reality of causality, and no new constructs were inserted into the new model (i.e., the PIFs and MCFs are the same between the two models), there was little new construction required to produce the new model. However, the cross-MCF dependencies had to be re-encoded because the sequencing of MCFs changed between networks. As a result, the situational and common context dependencies in the Fire Response model had to be recreated in the new network.

6.7.3 Recreate Cross-MCF Dependencies

Although submodels can be copied across GeNIe models, the relationships between submodels (e.g., situational and common context dependencies) will not be preserved in the new model. As a result, it is critical to review the new model for any dependencies that need to be re-built. For instance, if there are common context PIFs present, they will need to be reconnected to any cognitive phase nodes that are susceptible to common cause dependency. For instance, the I-, A-, and C-phase MCF nodes copied from the ISLOCA scenario had to have the Overarching Issues Workload PIF added as a parent node in the new model for the Fire Response scenario. Similarly, the situational dependencies have to be re-built to connect subsequent MCFs (if applicable). Because the cognitive phase patterns will change between network, care should be taken at this step to ensure that appropriate situational dependency relationships are included, as in the creation of the initial model (e.g., Chapter 5).

6.7.4 Verify Model Structure

The final step in creating the new model from existing idioms and submodels is, naturally, to verify the constructed model structure to ensure it is sensible and appropriate. For HRA, this includes ensuring that the nodes present in the model are all applicable to the scenario and that all appropriate dependencies have been encoded between the requisite nodes. For BNs, this also requires ensuring that no cycles or feedback loops were accidentally created in the new model, although most BN softwares will not allow this to occur in the first place.

6.7.5 Copy CPTs from Quantified Idiom Structures

Because the cross-MCF dependencies (i.e., situational and common context idioms) had to be recreated in the new model, the CPTs of the affected nodes were also recreated. Generally, there are two possible pathways available for quantifying the CPTs of nodes in a model created from existing idiom structures and BN fragments, as illustrated by Figure 6.5.

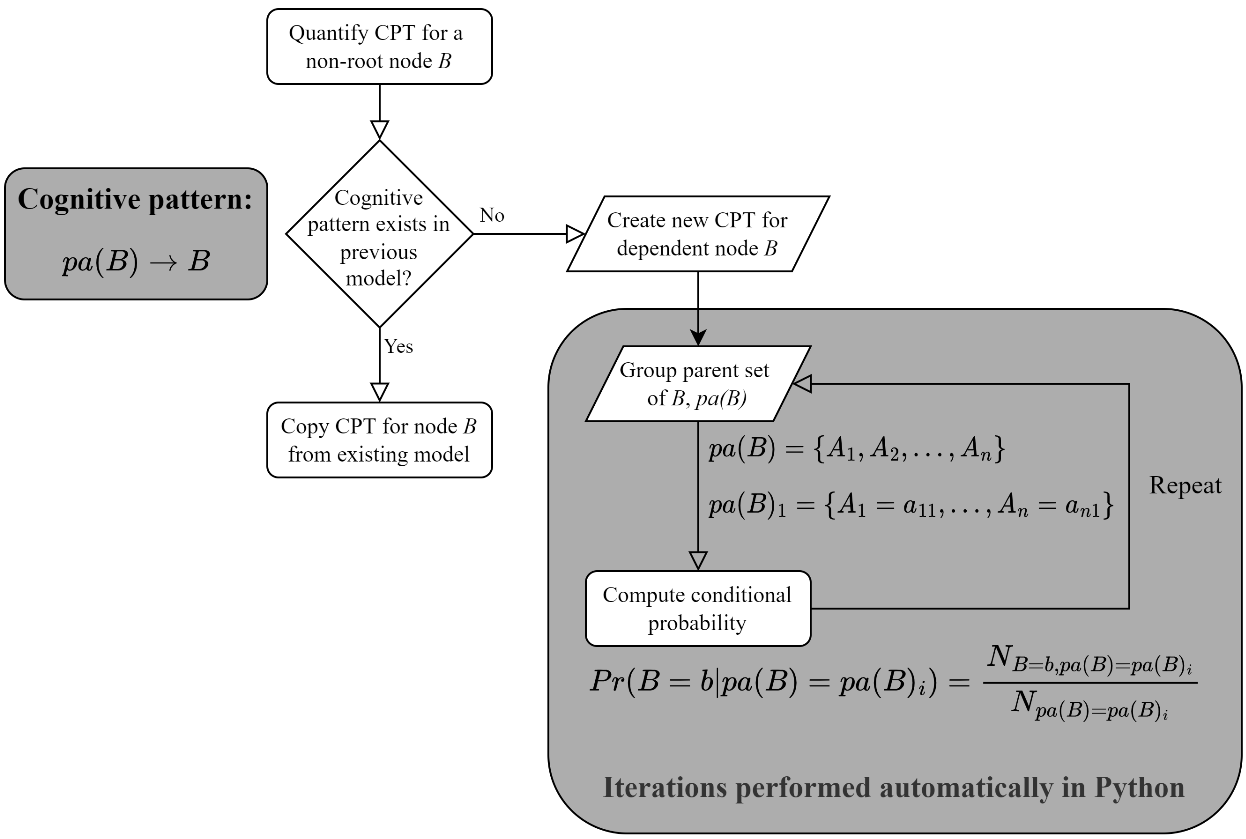


Figure 6.5: Quantification methodology for new model CPTs from existing idioms and BN fragments.

If one is attempting to quantify a non-root node (i.e., a node with at least one parent) in a new model built with idioms and BN submodels from an existing model, the first step is to determine whether the cognitive pattern observed in the new model, i.e., $pa(B) \rightarrow B$, is present in the existing model. For example, the cognitive pattern *A-phase MCF*$\rightarrow$*A-phase MCF* exists in the ISLOCA model and the A-phase PIFs are the same between the two models, therefore the CPTs for any node in an A-phase submodel that follows an A-phase submodel could be copied between the two models.

However, if the cognitive pattern observed in the new model is *not* present in an existing model, the CPTs for any node in a situational dependency idiom must be re-quantified with the parent set of the node in the new model. For example, the cognitive pattern *A-phase MCF*$\rightarrow$*C-phase MCF* must be quantified in the Fire Response model, but there is no $A \rightarrow C$ cognitive pattern in the existing ISLOCA model. As a result, the CPTs for C-phase nodes that are affected by situational dependency from the previous MCF must be re-quantified, i.e., the Time Criticality node in the C-phase MCF must be re-quantified in the Fire Response model. Once the re-quantification and/or CPT copying is complete, the model for the new scenario is fully quantified and ready for use in prognostic and/or diagnostic applications.

The process shown in Figure 6.5 can be used to requantify any node as required. For example, if new PIFs are identified for the new model and must be added to a given idiom, the states of the new PIF are simply added to the parent node set of the dependent node. The empirical conditional probability of each dependent node state is then computed as normal, and the CPT is created to include the newly-added constructs. Of course, Figure 6.5 applies to all types of data, including expert elicitation.

The idioms can therefore facilitate the creation of multiple scenario models by providing pre-built causal structures. This research shows that, when scenarios are modeled at the same or similar level of abstraction, the idioms and resulting BN fragments (with the associated CPTs) can be copied between models. Temporal dependencies (common context and situational dependencies) must be reconstructed, and the affected CPTs re-quantified. However, even if significant re-quantification is required, the creation of subsequent models from existing idioms and BN fragments proceeds much faster than the initial model creation and quantification processes described in Chapter 5 and the earlier sections in this chapter, especially when using pre-cleaned data (e.g., data used to quantify previous models) to re-quantify the CPTs.

6.8 Conclusions

This chapter created a methodology for quantifying HRA BN models built from the HRA dependency idioms. Employing the HRA dependency idioms when building and quantifying a causal BN facilitates the construction of BN models for full scenarios. Furthermore, this structures the quantification process such that it is amenable to both HRA data and more traceable expert elicitation. This methodology for HRA dependency quantification represents a significant improvement in quantitative HRA modeling by eliminating the use of arbitrary dependency levels and equations. By combining causal logic and data, this methodology improves our scientific understanding of dependency in HRA and provides, for the first time, a method to soundly model the effects of dependency in a formative way. This methodology is high-fidelity and creates fully model- and data-driven HRA BN models that are objective, traceable, and sound. Therefore, the quantification methodology developed in this chapter is more objective, traceable, and sound than previous HRA dependency quantification efforts.

The quantification methodology created in this chapter is suitable for any HRA model built from the HRA dependency idioms. However, quantifying dependency depends on the underlying logic of the specific HRA dependency idiom (i.e., definitional or causal). For definitional dependencies, quantifying the dependency requires understanding the connection between parent and child states and then applying a deterministic mapping between them. For instance, the observation of a CFM necessitates that the associated MCF is failed, because the CFM defines the failure of the MCF. Accordingly, the MCF will take state "Failed" when any associated CFM is "Observed," and "Successful" otherwise. For causally-oriented dependencies, e.g., causal, common context, situational, and effect-modulating, dependency quantification requires iteratively computing the conditional probability of each child node state, when conditioned on all the possible combinations of parent node states. In robust datasets, the required conditional probabilities are easily found in the data. For sparse or incomplete data, a combination of iterative sampling, expert elicitation, and the EM algorithm can be used to fill in missing conditional probabilities in a traceable and objective manner. In either case, the quantification process is streamlined by the HRA dependency idiom-based construction of the model, which identifies the conditional probabilities that are required, which thus optimizes the effort required to parameterize the network by preventing unnecessary calculations.

This chapter further showed that the HRA dependency idioms, as generalized causal structures, can be reused across scenarios like building blocks. If the same constructs (e.g., PIFs, CFMs, MCFs) are applicable then the pre-constructed BN submodels can be reused directly, and lower the time and effort required to create and quantify subsequent models. This presents an avenue for the quick construction of additional models with significantly reduced effort. Future models can be built quickly from the idioms and quantified fragments.

Chapter 7: Leveraging this Research to Improve HRA and Create an Effective Data and Modeling Lifecycle

This chapter provides a set of comprehensive recommendations designed to promote the inclusion of this research broadly within model-based HRA, and to improve the collection, storage and use of HRA data.

7.1 Introduction

For HRA, or any scientific field, data is what facilitates the development, revision, and application of new theories to solve real-world problems. Data and models have a symbiotic relationship, wherein data enables the creation of robust models which can in turn explicate areas of data needs and even be used to create new data. However, the history of HRA shows that this symbiotic relationship can be disrupted by the creation and implementation of disparate methods and data sources. This research created a unified definitional basis for HRA (Chapter 3), defined a set of HRA dependency idioms and their BN implementations (Chapter 4), and the methodologies for creating and quantifying BNs based on HRA data (Chapters 5 – 6). Accordingly, the results of this research can be used to improve both HRA modeling and the HRA data lifecycle, and reinforce the symbiotic relationship between data and models.

For engineering applications, data allows analysts to produce models that approximate the real scenario with appropriate fidelity. As a result, it is important that any available data be treated appropriately as long as it is in use, and that models are made to put the data to use. Building more robust models can improve our fundamental knowledge of human-machine team reliability, provide diagnostic and prognostic capabilities for real-world critical systems, and elucidate aspects of HRA theory that need to be further explored. Data, especially sparse data, can only provide so much to a field. Models allow analysts to glean additional causal insights for HMT reliability from the available data; these causal insights are critical to improving HMT reliability, and cannot be identified in data alone. Accordingly, the modeling constructs and methodologies created in this dissertation allow analysts to build HRA models that provide increased insight, improve HMT reliability, and ultimately guide the collection of more data.

Good models rely on good data – hence the adage that models can become beholden to the idea of "garbage in, garbage out." This may be especially problematic for HRA, which is a field inherently prone to generating sparse data. The inherent sparsity of HRA data can be corrected with quantification techniques such as the EM algorithm. However, it is essential that additional sparsity and other issues are minimized in data that is disseminated to analysts. This means that there must be protective actions taken to ensure the quality and quantity of HRA data is amenable to HRA modeling efforts. Data can be viewed as following a lifecycle that begins with planning the data collection scheme and ends with the publication and distribution of results [123]. The data lifecycle for HRA consists of three major phases, namely collection, storage, and use, during which it is manipulated by multiple stakeholders and analysts. Data lifecycle management typically refers to security, sharing, and use protocols for ensuring the appropriate use of data, which may be driven by a regulatory framework. The recommendations developed in the course of this research are not aimed at regulatory compliance but at ensuring the utility of HRA data for current and future applications. These recommendations form the first steps towards building a useful framework for collecting, storing, and using HRA data.

This chapter provides recommendations to the HRA community to leverage the results of this dissertation to improve both HRA modeling and the collection, storage, and use of HRA data. Continuing to improve the theoretical basis and modeling processes in HRA will help to build a more robust understanding of why, how, and how often human error occurs in complex engineering systems, and provide actionable insights for improving human (and therefore system) reliability.

The definitions and HRA dependency idioms created through this dissertation provide the building blocks to build more objective, traceable, and accurate causal models for HRA, which can leverage the higher quality data produced by using these recommendations to improve our fundamental knowledge of human reliability.

7.2 Improving the Theoretical Basis of HRA

The first objective of this research (Chapter 3) defined many useful concepts for the HRA context, including dependency, that will shape the way HRA is discussed, modeled, and quantified in the future. The definitional basis created in Chapter 3 standardizes the terminology and concepts that form the foundation of HRA, and as such provides a useful starting point for reversing the effect of over fifty years of siloization, that has been both a cause and effect of pursuing disparate models and data collection projects. The first recommendation for incorporating the results of this dissertation within HRA is, therefore, to use the definitional basis established herein to review and revise existing taxonomies of HRA constructs. This process can, for instance, work to "harmonize" different HRA methods with each other, so that results are more easily communicated and data sources can be more easily combined.

The terms and constructs defined in this research should be used, external to any HRA method, to evaluate existing taxonomies of the HRA constructs (e.g., the PIF taxonomy in [42] and the MCF taxonomy in [82]) to ensure that the constructs in these taxonomies align with the definition in this research. This will ensure that the foundations of HRA are consistent across methods, and thus that methods and data collection projects are amenable with each other. These taxonomies represent the basic HRA constructs that form various HRA methods and are used to communicate the insights gleaned about HMT reliability. As a result, it is critical that there exists a single, robust conceptualization of these constructs in HRA, which can be enforced by the inclusion of the results of this dissertation.

Using the terminology and causal constructs established in this research can be done without drastically altering the processes and methodologies in the various extant HRA methods. The practitioners of current HRA methods (e.g., THERP, ATHEANA, SPAR-H, etc.) should ensure that the constructs and terminology in their chosen method align with the results of Chapter 3. In some cases, the differences may be more semantic than technical (e.g., the PIF versus PSF), in which case redefining the construct in accordance with this research will not require changes to the methodology. In other cases, however, redefining a construct may result in changes to the qualitative or quantitative processes. For example, redefining the HFE in current HRA methods may require analysts to change the methods used to identify the HFE and/or quantify the HEP. However, such changes would improve the objectivity, traceability, and fidelity of the resulting model. The creators of emerging HRA methods (e.g., Phoenix, IDHEAS) and those who will embark on such missions in the future should similarly ensure that the constructs included in their methods align with the definitions provided in this research. This will entail incorporating the modeling constructs created in Section 3.4.1 (visualized on a BN structure in Figure 4.2). Further, new methods should strive to both incorporate and inform HRA data to the greatest extent possible, to maintain the symbiotic relationship between HRA data and models.

Finally, and perhaps most saliently for this research, the definition of dependency (Section 3.6.3) and idiom framework (Chapter 4) be used to conceptualize and model dependency between HRA constructs. Incorporating this definition as the standard in HRA methods and guidance documents (e.g., NUREGs and ANS/ASME standards) will ensure that HRA works from a consistent understanding of what dependency means; a definition that is rooted in probabilistic causality and applicable between multiple HRA constructs. This also ensures that HRA models model and quantify dependency in a traceable, objective, and sound manner that supports cross-method validation.

Improving the theoretical basis of HRA, both in general and regarding dependency, will allow HRA to rescue the *science* from the current state of the field, which is closer to a loose collection of terms, concepts and methods than to a defined science. Having a uniform and well-understood basis will support additional studies into psychological, physiological, and engineering aspects of HRA, and facilitate the creation of a normative theory of HRA. This research provided many well-defined constructs and a robust understanding of dependency, to set up a normative theory to understand what constitutes reliable HMT performance, which will be further informed by sound HRA modeling applications.

7.3 Improving HRA Modeling

Beyond the definitional basis for HRA created in Chapter 3, this research also created the set of HRA dependency idioms and methodologies for the construction and quantification of HRA BN models. These represent foundational improvements for the conceptualization, modeling, and quantification of HRA, and should be leveraged to create robust HRA models. Including the methodologies developed in this dissertation in both mature (e.g., ATHEANA, SPAR-H) and emerging (e.g., Phoenix, IDHEAS) HRA models can facilitate the modeling and quantification of formative dependency-inclusive models of HMT performance. The modeling constructs and methodologies can be leveraged to improve HRA modeling in several ways, including but not limited to the following:

1. Trace, validate, and improve existing HRA methods.
2. Use the idioms to guide causal thinking.
3. Create HRA methods incorporating BN idioms.

7.3.1 Trace, Validate, and Improve Existing HRA Methods

The most immediate application for using the idioms is to use them as a framework for tracing, validating, and ultimately improving *existing* HRA methods. For HRA methods that incorporate BNs (e.g., Phoenix [82] and IDHEAS [45]), the idioms can be used to trace existing BNs and ensure that the dependencies included are sensible, validate the quantification of the BNs, and build improved BNs for use in the HRA method. For non-BN based HRA methods (e.g., THERP [16], ATHEANA [48], SPAR-H [31], etc.), the idioms can be used to validate and/or improve the quantification of dependency by using the values obtained from a quantified idiom-based BN and/or substituting the BN for the checklists and equations used in such methods.

For BN-based methods, the idioms should be used as a framework to assess the structure of the networks. Evaluating the BNs through the idioms entails determining whether the included dependencies are understandable and quantified appropriately. Understandable dependencies are those that make sense within the idiom structure, are supported by causal or definitional logic, and connect the appropriate variables (within the semi-hierarchical BN structure). The idioms will also show whether the dependencies are quantified appropriately (i.e., the probabilistic or deterministic evaluations of conditional probability are correct), and be used to re-quantify the network if necessary. The idiom framework will support the traceability of the overall model, by providing the framework to provide documentation of the existing dependencies and a framework for documenting added dependencies. Finally, the idioms can aid in communicating the impact of the dependencies to analysts and stakeholders.

Applying the idiom framework to non-BN HRA methods can update their dependency assessment methodologies in light of the results produced in this research. In many ways, this would be a complete overhaul of the dependency assessment methods, from the conceptualization of dependency as a concept to the quantification of conditional probabilities. In non-BN HRA methods, the assessment of dependency is typically done with checklists and/or decision tree structures. These can be transformed into BN structures that represent each outcome of the decision tree (or each possibility of the checklist), which can then be compared to the idiom-built BNs. The idiom structures can replace the checklists or decision trees with a traceable, supportable set of dependency relationships that are rooted in causality and validated with data.

As demonstrated in this research, it is possible to build HRA BNs (Chapter 5) from the idioms and available HRA data, and use the resulting network to assess the performance of HMTs in given scenarios (Chapter 6). This represents the best opportunity to impact the practice of HRA. Replacing existing HRA processes with the idiom-based processes developed in this dissertation will transform HRA dependency from a subjective, unsound, and untraceable process into a rigorous, traceable, and objective enterprise rooted in causality. BN representations of existing HRA processes have been constructed for, e.g., SPAR-H ([63, 65]), but these did not substantively change the assessment of dependency. Rigorous implementations of existing HRA methods with BNs should use the idioms to guide the model construction and enforce dependency.

7.3.2 Use Idioms to Guide Causal Thinking

In addition to structuring causal models in HRA, the idioms provide the structure to guide the process of causal inquiry. The idioms, as fundamental logical relationships, modularize the causal web upon which HRA is woven to provide understandable relationships with associated graphical structures. This allows analysts to investigate causality in a piece-wise manner and iteratively build to larger, full-scope networks. By taking a pair-wise comparison of HRA constructs, within the bounds set by the semi-hierarchical BN structure for HRA (e.g., relationships traverse no more than one level of the hierarchy), and applying the methodology depicted in Figure 5.5, analysts can iteratively build the dependency structure that connects the nodes in the model.

The idioms defined in this research (Chapter 4) should be used to guide the creation of causal models for both existing and novel HRA applications; while the specific PIFs, CFMs, and MCFs may change with application and context, the structures linking these constructs represent the reality of causality, rather than construct-specific relationships. Applicable uses of the idioms extend beyond building causal models to developing more effective data collection protocols and understanding the relationships engendered by the use of procedures, guided by idiom-based causal logic.

While the idioms, and resulting BN structures, can be used to augment existing HRA methods, they can also be used to entirely replace existing HRA methods and guide causal thinking when incorporated into PRAs. The input from HRA to a PRA is typically the HEP, the "final" probability of failure obtained from the network that governs the branching probabilities of a basic event in the PRA. This value, of course, can be generated in a variety of ways ranging from expert elicitation to quantitative HRA methods. Therefore, one pathway to incorporating the idiom-based methodology is to build idiom-based BN representations of PRA basic events, to act analogously to the fault trees that are used to find the probability of hardware failures. The idioms will not only facilitate the robust representation of single basic events (i.e., MCFs) with dependencies included between PIFs and CFMs, but also allow for cross-MCF dependency to be accounted for by linking the BN fragments for multiple basic events.

In addition to building the model for individual MCFs as basic events (Figures 5.1 – 5.4), the cross-MCF dependency idioms can be used to explicate the structure of the event tree where human events are concerned. For example, if situational dependency is expected between MCF A and MCF B, the portion of the event tree containing these MCFs will have to

The idioms created in this research (Chapter 4) facilitate building both BN submodels for individual MCFs as basic events (Figures 5.1 – 5.4) and for larger models with basic MCFs by enforcing cross-MCF dependencies (e.g., common context, situational). Using BNs to represent basic events in a PRA ensures the traceability, objectivity, and realism of the HRA quantification, and therefore the parent PRA itself.

7.3.3 Build HRA Models from the Idioms

The HRA dependency idioms, and the construction and quantification methodologies, developed in this research provide all the tools required to build full-scope HRA models for multiple applications. While they can be used to immediately replace the dependency-specific aspects of existing HRA methods, as discussed above, additional model fidelity can be gained by developing new models with the HRA dependency idiom framework as a basis. This is a particularly salient trait of the dissertation results, as the nature and context of human performance in complex systems is changing with the advent of increased automation and novel paradigms of operation (e.g., digitized, remote, ex-control room operations). These developments are catalyzing a review and revision of the traditional HRA methods rooted in behaviorist psychological principles applied within static, controlled environments. As a result, there is a need to apply the strong, causally-oriented and context-agnostic foundations created in this research to promote the development and deployment of HRA methods that will be amenable to the modern context of HMT performance.

The HRA dependency idioms, as representative models of context-aspecific causal relationships between HRA constructs, represent a good opportunity to build HRA models from the ground-up to work within these non-traditional contexts. For example, the idioms could be used to conceptualize, model, and quantify HMT performance for the response to external hazards at nuclear power plants (e.g., high winds, flooding, seismic hazards). Current HRA methods may not be broadly applicable to modeling such contexts ([121, 124]), and as a result there is a need to build causal models from first principles, for which the idioms can provide both the analytical and modeling structure. The results of this dissertation will, therefore, be implemented to increase our understanding and ability to model human reliability under novel contexts including external hazards.

Beyond external factors, the response to which is critical to understand for a robust HRA of critical infrastructure, there are also concerns regarding the applicability of existing HRA to the increasingly digitized and even automated internal environments in which HMTs operate. This is coupled with a changing dynamic between the components of the HMTs, in which the humans are increasingly being transitioned to a more advisory or oversight-focused role. These changes are incompatible with many existing HRA methods, and potentially antithetical to current conceptualizations of HRA. However, the definitions and idiom framework created in this dissertation could facilitate robust HRA modeling of these new contexts. Because the HRA constructs defined in Section 3.4.1 relate to the performance of the HMT, rather than the humans specifically, the changing dynamic between human and technology does not affect the validity of the constructs. Further, the idioms represent the causal structure between the HRA constructs in general, rather than specific human tasks/functions. This means that the idioms are amenable to modeling a scenario regardless of the dynamic between the components of the HMT.

The idioms therefore provide the ability to model HRA in novel contexts and applications, which is of critical concern in light of the potential conflicts between existing HRA and the new paradigms of HMT performance. The idioms also provide the advantage of structuring the model building and quantification, through the methodologies developed in Chapters 5 and 6. This will provide the ability to quantify the resultant model from applicable aspects of available data, as well as indicating the portions of the model that will require new data to be collected. This research therefore provides the building blocks and methodologies for constructing robust HRA models for novel contexts and applications.

While the technical contributions of this research provide many capabilities for improving HRA, realizing the promised gains will require a higher quantity and quality of data than is currently available in most HRA sources. Accordingly, this research identified areas for improvement within the HRA data lifecycle that, if incorporated, will provide the quality data that is required to engender improvements in the conceptualization, modeling, and quantification of HRA.

7.4 Improving HRA Data Collection

Figure 7.1 below (adapted from [123]) illustrates a high-level data lifecycle for HRA. The collection phase includes planning the data collection and management scheme, as well as the actual acquisition and any pre-processing that is performed before storage. Data storage entails reviewing, updating, and sharing the data with other analysts, so that they can finally use the data by building, quantifying, and publishing models. Each phase of the data lifecycle is accompanied by administrative processes that provide documentation, curation, and security for the data. The sections below provide recommendations at each phase of the data lifecycle that will ensure the utility and quality of HRA data for future research, both related to this work and genearlly within HRA. These recommendations are not meant to simply improve HRA data for the modeling and quantification of dependency. They will accomplish that goal, but also ensure that HRA data is amenable to a variety of uses. Similarly, these recommendations are not meant as a call to entirely discard or discreted the existing data collection schemes, but as a guide to constructing enhanced HRA data collection projects that can augment the supply of available data.

One of the main stakeholding and contributing fields to the collection and use of HRA data is nuclear power operations. As a result, there are multiple ongoing HRA data collection projects with application to nuclear power plant operations, e.g. SACADA and HuREx ([84, 102]), which are thoroughly compared in [103, 104]. These projects collect data under distinct guiding method-

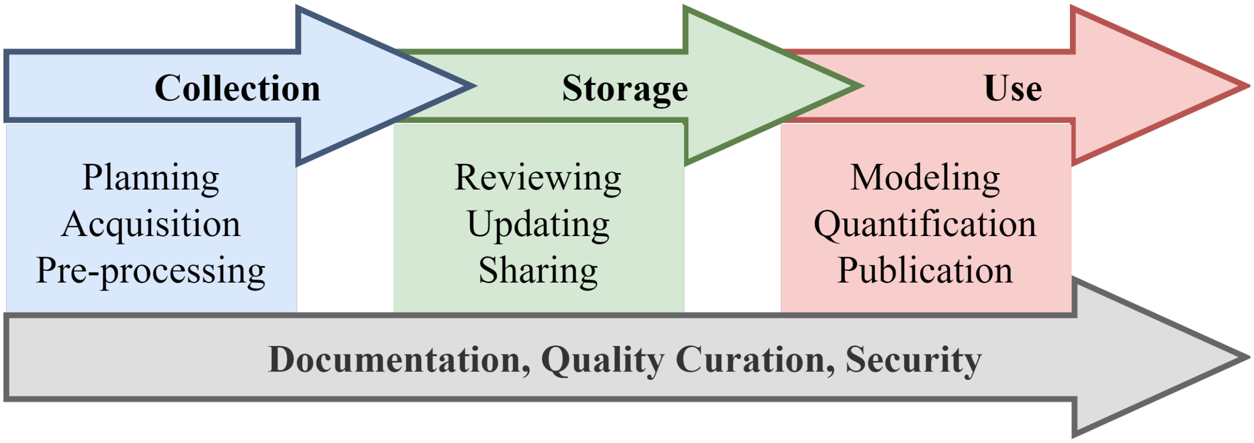


Figure 7.1: High-level HRA data lifecycle.

ologies and use their own taxonomies, but both facilitate data collection during full-scope simulator training of nuclear reactor crews [103]. Collecting data during routine operator training ensures that HRA data is collected in realistic scenarios and the time, money, and personnel costs are not prohibitive for either the stakeholders or researchers. However, the collection of HRA data is susceptible to pragmatic and idealistic flaws which limit the utility of the data for HRA purposes. These flaws can be mitigated generally by a consideration of the following recommendations, explained in the subsequent subsections:

1. Record narratives and crew comments for each scenario.
2. Use existing HRA causal constructs.

7.4.1 Record Corresponding Narratives

One of the critical aspects of causal modeling is being able to piece together the causal structure from data, including identifying the constructs that serve as causes to temporal dependencies, e.g. situational dependency. Numerical data is not the only resource available to build a causal structure. Narrative data (e.g., crew/observer comments, procedures, narrated videos, event reports etc.) can be used to identify the causal structure of a scenario [65]. Narratives can serve an important role in illuminating the causal structure by providing analysts with additional evidence of the causal role of HRA constructs. For instance, numerical data may reveal that there is generally less time available for action tasks following a misdiagnosis, and therefore a higher human error probability during such action tasks. The numbers will not communicate the causal mechanisms for this relationship, but insights into the causal mechanisms can be gleaned from narrative.

Collecting (and providing) corresponding narrative accounts of the simulator data is essential to ensuring the traceability and realism of the resulting model, but also prudent practice to maximize the impact of simulator studies for HRA. Narrative can contain useful information on causal structure and mechanisms, but can also serve as a guidepost to cleaning and updating the numerical data, adding additional factors that were not included in the numerical data, and improving analyst understanding of the scenario progression. Basic narratives and crew comments are available for some SACADA scenarios, but this is not uniformly true.

One of the most readily-implementable ways to improve HRA data collection is therefore to collect and disseminate more extensive and detailed narrative information. This can be done in numerous ways, not all of which are feasible in the context of collecting data from high-fidelity full-scope simulators. Narrative information can be collected by providing the procedures used, interviewing crews after the simulation, videotaping the simulator, and/or having observers develop an annotated timeline of the events. Procedures, where available, can provide context to *why* the crew may have made seemingly-incorrect choices. Crew interviews and comments should be obtained for the scenario and disseminated as part of the data; the crew can provide insights into unobservable PIFs (e.g., stress, mental models, etc.) that can help expand and validate the causal structure of the model. Videotapes or annotated timelines can expand on the environment, crew cohesion, and organizational factors present in the simulation.

Providing narrative information along with numerical data can greatly improve the usability of the data, and the traceability of resulting models. Collecting the narrative information will not require significantly more effort from observers, as procedures are ubquitous in the simulator environment, crew interviews and debriefs are standard practice, and videotaping is becoming more common [103, 104]. Data sharing and security protocols may limit the ability to disseminate some of these artifacts (particularly videotapes and procedures). However, providing the crew interview/debrief transcripts and observer notes will serve as an important step towards the greater inclusion of narrative as part of a cohesive HRA data package.

7.4.2 Use Existing HRA Causal Constructs

HRA data collection projects (e.g., SACADA and HuREx) often use distinct, and sometimes incompatible, taxonomies of constructs when collecting data. This limits the utility of the data for existing and emerging HRA methods, which can require translating the data between taxonomies or discarding data that cannot be translated. Further, this limits the compatibility of different datasets with each other, potentially bottle-necking the supply of data. This partitioning serves a purpose in facilitating research at multiple levels of abstraction, such as sub-task level and function-level modeling, but may also be part-and-parcel with the tendency to create new, distinct HRA methods rather than improve the theoretical foundations of existing methods. However, there is a pathway to both harmonizing the disparate HRA data streams *and* maintaining the multi-level modeling capabilities afforded by the different approaches. The set of causal constructs defined in Chapter 3 provides the taxonomy and hierarchy to allow for a unified data stream that can be used for multi-level modeling.

In this recommendation, a unified data stream consists of multiple independent data collection programs that each collect a subset S of modeling constructs pulled from a standard set A, $S \subseteq A$ of available modeling constructs (based on the constructs defined in Chapter 3). For instance, a group in the USA could collect 10 PIFs, 7 MCFs, and 15 CFMs, while a second group in South Korea collects 15 PIFs, 5 MCFs, and 12 CFMs. These projects will not collect identical datasets – the resultant datasets may even be mutually exclusive – but they will be *compatible* with each other because of the shared theoretical basis. These datasets could then be combined with minimal translation effort to form a larger and/or more detailed dataset.

In addition to unifying the data stream, using existing causal constructs will ensure that the data collected is applicable to causal modeling applications. Adhering to the definitions of the causal constructs provided in Chapter 3 will ensure that the PIFs collected are well-suited to the needs of

HRA analysts and capture the relevant aspects of the environment, organization, system interfaces, crew composition, task characteristics, and operator personalities that will influence HMT performance. The data used in this research consisted mainly of PIFs relevant to task characteristics (see Table 5.2). Thus, there is a significant opportunity to improve the applicability of this research with additional data, by collecting more environment and/or crew-oriented PIFs in addition to the task characteristics. Additionally, following the idiom structure will allow simulator studies to collect information on specific CFMs and further inform causal chains, and common contexts important to HRA, such as the causal chain through Available Time and the common context of Workload, so that these causal structures can be quantified with higher fidelity.

Standardizing the basis for data collection does not, alone, facilitate multi-level modeling. That is engendered by, where possible, collecting data at *lower* levels of abstraction, e.g., erring on the side of low, rather than high, levels of abstraction. This can be done in addition to, or replacement of, collection at higher levels of abstraction. For performance constructs, this means collecting data on sub-tasks as well as (or instead of) tasks and functions; For PIFs, this entails collecting fundamental PIFs (e.g., the lower levels of the PIF taxonomy defined by Groth and Mosleh [42]). High-level data is not easily broken into lower-level constructs, which requires either additional information or data to quantify the fragments. However, definitional or construction (with a well-documented quantification logic) dependency can be used to combine low-level data into high-level modeling constructs. For example, observed fundamental PIFs can be combined to form a higher-order PIF (e.g., Task Load and Non-Task Load form Workload) if the analysts develop a suitable quantification architecture to do so (e.g., fuzzy logic, noisy-Adder, etc.), but the reverse cannot be done without making significant assumptions about the unobserved states of the fundamental PIFs.

7.5 Improving HRA Data Storage

The storage of HRA data is tantamount, if not paramount, to ensuring the appropriate collection of data. Collected data should not be stored away for posterity, but should be tabulated, shared, and implemented to ensure that analysts can better understand, and thus effectively reduce, the contribution of humans to system risk. As a result, storing HRA data in a manner that is amenable to constructing and evaluating causal models of human reliability is critical to ensuring the continued improvement of the field. Building and quantifying BNs from data requires being able to identify the constructs/variables recorded, the possible states each variable can take, and the state of each construct at each recorded point (for HRA, each MCF/task performed). This requires formatting the stored data tabularly, to facilitate identifying each of these characteristics within the dataset. This is how SACADA and most engineering datasets are provided, but there are ways to improve on current capabilities. For this application, and future research, the best way to facilitate the building and quantification processes may be to store each simulated scenario as a separate record, instead of grouping different crews together because they performed the same scenario. This would make the data processing for BN construction easier by removing the data expansion step. Additionally, this would allow the observers to store more cognitively-focused factors that will vary between crews, and thus could be missed if different crew records are combined. The following recommendations should also be considered when developing a data storage paradigm to ensure that the data is useful, contemporary, and complete:

1. Ensure that data is as complete as practicable.
2. Regularly review, update, and disseminate stored data.
3. Tabulate summary statistics as secondary outcomes, rather than primary targets.

Building and quantifying BNs from data requires being able to identify the constructs/variables recorded, the possible states each variable can take, and the state of each construct at each recorded point (for HRA, each MCF/task performed). This requires formatting the stored data tabularly, to facilitate identifying each of these characteristics within the dataset. This is how SACADA and most engineering datasets are provided, but there are ways to improve on current capabilities. For this application, and future research, the best way to facilitate the building and quantification processes may be to store each simulated scenario as a separate record, instead of grouping different crews together because they performed the same scenario. This would make the data processing

for BN construction easier by removing the data expansion step. Additionally, this would allow the observers to store more cognitively-focused factors that will vary between crews, and thus could be missed if different crew records are combined.

7.5.1 Ensure Data Completeness

Human reliability data is inherently sparse. With typical marginal HEPs on the order of 1-2%, thre will be heavy class imbalance between the “Success” and “Failure” states of the MCF data. When coupled with the number of factors that may be recorded in the dataset, tens of thousands of simulated MCFs may be required to obtain a reasonably complete dataset. The problem of sparsity can be corrected post-collection by implementing reparameterization algorithms (e.g., EM) and/or using structured expert elicitation processes (based on the HRA dependency idioms) to fill in some of the missing data. However, these methods correct for a problem that is inherent in HRA (sparsity), and which need not be exacerbated by additional missing data. Experimental design can also be used to identify missing combinations of factors in the present data and design new scenarios for crews to run that will contain those factors.

Importantly, the sparsity problem can be exacerbated by other sources of incompleteness in data, e.g., human error in collection and storage, and/or corruption during storage. Data collection agencies should make every effort to ensure that the data collected is reasonably complete, and that the stored data is accurate to what was collected. This requires identifying and correcting instances of input error and corruption in the datasets prior to disseminating the datasets for use by researchers. If the researchers were not involved in collecting the data, they may be unable to meaningfully correct errors in the datasets without compromising the robustness/application of their methods.

As a result, one of the most important aspects involved with storing HRA data is to ensure that the data is reasonably complete, and that the stored data is free from errors and erroneously missing values. This will ensure that the process of analyzing the data to build causal models is not further complicated and that the data is well-safeguarded.

7.5.2 Review, Update, and Disseminate Data

Part of a robust data lifecycle management protocol is regularly reviewing and updating the stored data. This means that the available datasets, especially if they are publicly-hosted, should be *regularly* screened to identify and correct missing values and corrupted data. Additionally, this requires appending the existing dataset with new data collected since the last period of updating. Continually reviewing and updating the dataset will ensure that the dataset is current and continually growing, both of which are characteristics of a healthy dataset.

Review of the dataset should be informed by experts both within the data collection agency and by the experts that are using the data, who may be able to identify specific issues in the data not found by data collectors. A review (and subsequent updating) of the data should be performed regularly, although the specific interval may be open to interpretation. Data reviews should not only identify issues with the currently-available data, but also provide guidance on how future data collection projects can better meet the demands of the analysts.

Another important aspect of data is that they reflect the current capabilities, concerns, and trends of the industry. For example, if the nuclear industry is exploring the possibility of digitized control rooms, the dataset should be updated as studies are produced to reflect that context. Updating the data regularly also means that researchers can stay ahead of where the industry is moving.

7.5.3 Use Summary Statistics Sparingly

The principal outputs of HRA data collection and analysis activities should be a more rigorous understanding of human-machine team reliability and the raw data to support modeling, not simply an updated value to input into the parent PRA methodology. As a result, the summary or descriptive statistics of the HRA data (e.g., HEPs, average values, etc.) should not be the primary outcome of the data collection project. This means that all of the data that goes into developing these summary statistics is presented with minimal alteration from the raw values. Instead of providing computed values, such as functions of multiple input variables, datasets should mainly present the recorded states of fundamental, observable characteristics in the scenario. Summary statistics are useful outputs of data analysis, but should not become the focus of HRA data collection and storage, which should prioritize providing raw data that can be used in a variety of applications. Summary/descriptive statistics can always be computed from raw data, but the reverse is not true. As a result, every effort should be made to share raw data rather than the summary statistics.

7.6 Improving the Use of HRA Data

For analysts, the most tangible or salient phase in the data lifecycle may be its use in the construction, quantification, and application of models. However, the ability to do so rests on the underlying labor in the collection and storage phases. Using the data is reliant on a robust collection program and useable storage framework, but the use of data imposes its own requirements for maintaining the data lifecycle. Preserving the documentation, quality and security of the data, and using the resultant models to propel the field forward and glean new insights into the nature of human-machine team performance and reliability, are critical aspects of data lifecycle management for HRA. Implementing the following recommendations will allow data to be used in a manner conducive to improving the HRA field:

1. Use the appropriate data.
2. Build better understanding with models.
3. Communicate results to stakeholders.

7.6.1 Use Data Appropriately

As with any venture that bridges data to models, it is incumbent upon the analyst to ensure that data is used appropriately, i.e., that data is used only to inform applicable/relevant models. This means that analysts should ensure that their data source matches the proposed model, with respect to multiple facets such as level of abstraction, construct definition, and applicable context.

For example, the SACADA dataset ([84]) is not an appropriate match for subtask-level modeling efforts. This would have required analysts to decompose constructs into finer levels of detail, entailing making many unsupportable assumptions. Similarly, the HuREx dataset ([102]), which collects data at a lower level of abstraction, was not an appropriate match for this research building a function-level model; a *majority* of nodes would have required construction dependency and/or translation to arrive at higher-order constructs, and function-level data was available in the SACADA dataset. Although building up in levels of abstraction is more easily supported than decomposition, the availability of data at the desired level of abstraction (SACADA) rendered the use of lower-level data unnecessary. Analysts should ensure, therefore, that the data they incorporate is appropriate to the level of abstraction of their model, or that appropriate processes are in place to support translating between the data and model. Collecting data at multiple levels of abstraction can allow for broader uses of data in building models (particularly when paired with construction dependencies for synthesizing higher-order modeling constructs from lower-level data).

Analysts should ensure that the constructs as defined in the data source (e.g., PIFs, CFMs, MCFs) align with the constructs defined for their modeling application. For instance, the model created for this research used the SACADA SFs constructs directly as model PIFs. However, it is also possible to translate between data constructs and modeling constructs, as in this research where the SACADA TOEs were incorporated into the model as cognitive-phase MCFs. Translation should be well-documented and the quantification scheme (if using construction dependency) supported by literature, data, and/or expert elicitation. Researchers should align their modeling concepts with the causal constructs from Chapter 3 to facilitate the verification, validation, and broader use of their models.

HRA data has been generated for, and used in, a variety of contexts spanning from nuclear power main control rooms to plane cockpits to oil rig operations. However, HRA data collected for a specific context may not be broadly applicable (e.g., oil rig operations and plane cockpits are significantly different), and industries may have different priorities as to the type of constructs and data that should be collected. As a result, it is important for analysts to understand both the context in which the data was generated, as well as the context of the model to which the data is to be applied. Further, analysts should endeavor to match contexts between available data and their model, or sufficient significant justification as to why the data is usable for quantifying the model.

Using HRA data appropriately in models ensures that HRA data is well-utilized to model applicable/similar contexts and that HRA maintains its validity as a field. Additionally, the appropriate use of data assists with ensuring data quality. For data-collectors, ensuring data quality is related to ensuring the completeness, applicability, and accessibility of datasets. For data-users, ensuring data quality means faithfully representing the datasets in publications and models, citing the source of data, and abiding by standard research practices for the use of data.

7.6.2 Build Better Understanding

Data and models are not only useful for gaining causal insight into a single scenario or application. Data and models can, and should, be also used to investigate and refine the theoretical foundations underpinning multiple aspects of HRA. For example, this research was not aimed at producing a novel diagnostic or prognostic model for immediate application in HRA; rather it was directed at building the theoretical foundations for the conceptualization, modeling, and quantification of dependency in HRA. In similar fashion, future uses of HRA data should support similar research focused on theoretical improvements, building on the achievements put forward in this dissertation.

There are areas of critical need in HRA that can be informed by HRA data collected under the guidance of these recommendations, including establishing the foundations for dynamic HRA methods and external-hazards HRA, harmonizing time-based reliability and traditional HRA methodologies, and investigating higher-order causal reasoning paradigms to produce HRA insights. Each of these approaches will advance HRA in a multitude of ways, produce improvements to the theoretical foundations of HRA, and critically require applicable HRA data to perform. The results of this research – namely, the causal construct definitions (Chapter 3) and idiom structure (Chapter 4) – provide a common basis for these data-based investigations and ensure that the methods and results of each are inter-compatible and easily translatable. Having a common basis for disparate research areas allows their pursuit to drive the field in a unified direction, rather than forming distinct, irreconcilable branches.

7.6.3 Communicate Results

Perhaps the most impactful way in which researchers can further the field of HRA is to share the insights gleaned from their use of data with stakeholders, researchers, and the public. Communicating research is essential to ensuring that methodologies and results are sound, applicable, and useful. Sharing research shows the efficacy of the work for solving real problems in industry, informs the development of next-generation HRA practices and methods, and guides data collectors to tailor their processes to support the field. Sharing positive and negative findings from HRA modeling efforts can inform data collectors about rising concerns in the field, new constructs to collect, and/or new scenarios to study.

7.7 Conclusions

The definitions (Chapter 3), HRA dependency idioms (Chapter 4), and methodologies (Chapters 5, 6) created in this research can serve multiple purposes if incorporated broadly into HRA. Using the definitional basis created in this dissertation can build a unified theoretical basis and facilitate the creation of comprehensive HRA models, which can be built in a traceable and objective manner with the HRA dependency idioms and construction methodology. Using the HRA dependency idioms to guide the collection, storage, and use of HRA data will not only provide the field with a widely-implementable data set that captures the constructs relevant to HRA. This data can then be used to objectively and accurately quantify the causal model to gain a deeper causal insight into HMT reliability, as well as diagnostic and prognostic assessments of real scenarios.

The definitions, idioms, and methodologies created in this research should be used to provide a firm technical basis for HRA conceptualization, modeling, and quantification. This will promote the creation of a standard, normative theory for HRA and help the transformation from a collection of HRA methods into a scientific field of HRA. Further, having a standard theoretical, technical basis will ensure that HRA modeling and data collection efforts can intercommunicate and that results can be shared easily and used to further the science of HRA.

The continued collection, storage, and use of HRA data is essential to not only furthering modeling efforts, but also to investigating and furthering our understanding of the foundational concepts in HRA. Both aspects of improved HRA data will help to reduce the human contribution to system failure. Explicating the causal web of "human error" and producing more objective, accurate, and traceable models will allow analysts to produce more effective mitigating actions.

The recommendations provided in this chapter are guidelines to improving multiple facets of HRA, including the data lifecycle. However, these are not necessarily applicable to each and every data collection project, each of which will be designed with specific goals/aims in mind. However, the use of standardized constructs for HRA data, and recommendations on the storage of data, will ensure that collected data is amenable both to the stated purpose of the collection project and more broadly within HRA.

Chapter 8: Research Contributions, Impact, and Pathway Forward

This chapter provides an overview of the contributions and impact of this research, including publications, presentations and their impacts on the discipline of HRA and reliability engineering. Publications and presentations communicate the research process and results to HRA practitioners and researchers to provide a basis for continued work; these publications undergo peer-review, a cornerstone of the scientific process. This dissertation addresses some of the overarching limitations in HRA and closes key technical gaps related to dependency. This research has produced both theoretical and practical contributions to the field, and has set up a framework for research into the study of dependency and HRA going forward.

8.1 Research Summary

This research project pursued three main research objectives (ROs), each of which addressed a technical gap in HRA and produced significant contributions for the conceptualization, modeling, and quantification of HRA dependency. The principal results of each research objective are summarized below.

Research Objective 1: Define the lexicographical and mathematical foundations of dependency in HRA through the standardization of definitions and mathematics for core dependency concepts. RO-1 explicitly addressed the lack of a unified lexicographical and mathematical dictionary for HRA. As a result, the research provided a cohesive set of core definitions for foundational terms and concepts that HRA currently lacks.

8.1.1 Key Results: Definitions (RO-1)

The results of Research Objective 1 provided definitions for common and foundational terms in HRA. The full definitions, created in Chapter 3, are summarized below:

HRA Dependency:

A dependency relationship exists between two HRA variables if they are connected by a direct or indirect causal relationship that changes the conditional probabilities of the variables; dependency exists regardless of whether the existence or utility of the variables is acknowledged within HRA.

HRA Constructs/Variables (see Table 3.1):

- *Human Machine Team* (HMT): The combination of humans and systems that work symbiotically to perform functions and accomplish objectives necessary to maintain system safety and/or operation (see Section 3.4.1.1).
- *Human Failure Events* (HFEs): The highest-level failures considered in an HRA; the failure to complete a high-level objective (see Section 3.4.1.6).
- *Human error probability* (HEP): The probability of experiencing an HFE during a scenario (Section 3.4.1.7).
- *Major Crew Functions* (MCFs): The high-level actions of the HMT in a scenario (see Section 3.4.1.5).
- *Crew Failure Modes* (CFMs): The failure pathways that define *how* a function was failed (see Section 3.4.1.4).
- *Crew Activity Primitives* (CAPs): The lowest-level of abstraction actions/decisions undertaking by HMTs to achieve a function (see Section 3.4.1.3).
- *Performance Influencing Factors* (PIFs): The set of environmental, organizational, system, crew, and personal factors that describe the context in which the HMT operates and affect performance (see Section 3.4.1.2).

Task Decomposition Terminology (see Table 3.2):

- *[System] Event*: A milestone or system change that represents a significant moment in the scenario response or operation (see Section 3.4.2.1).
- *Objective*: The goal of the portion of event response under study, at the highest level of abstraction; used to discretize events (see Section 3.4.2.2).

- *Function*: A high-level system purpose that the HMT must perform to achieve an objective (see Section 3.4.2.3).
- *Task*: A fundamental action on the system taken by an HMT to perform a function (see Section 3.4.2.4).

Set Theoretic Concepts (see Table 3.3):

- *Union*: The group of elements belonging to any of the sets of interest (see Section 3.5.1.1).
- *Intersection*: The group of elements belonging to all of the sets of interest (see Section 3.5.1.2).
- *Mutual Exclusivity*: The condition wherein sets have no common elements, i.e., their intersection is empty (see Section 3.5.1.3).

Statistical Concepts (see Table 3.3):

- *Marginal Probability*: The probability of an event or variable occurring without considering other events in the universe (see Section 3.5.2.1).
- *Joint Probability*: The probability that multiple events or variables occur together (see Section 3.5.2.2).
- *Conditional Probability*: The probability of an event or variable occurring given evidence that another event or variable has already occurred (see Section 3.5.2.3).
- *Independence*: The condition wherein two or more events or variables have no connection, and thus have conditional probabilities equivalent to their marginal probabilities (see Section 3.5.2.4).

Research Objective 2: Create a causal framework of HRA dependency idioms that models the possible relationships between HRA constructs. RO-2 explicitly addressed the lack of an exhaustive and orthogonal set of dependency relationships in HRA. The resulting HRA dependency idioms created in this research effectively model archetypical logic structures and serve as the building blocks for a full causal picture of an HRA scenario.

8.1.2 Key Results: Idioms (RO-2)

The results of Research Objective 2 provided the set of archetypical relationships for HRA constructs, known as HRA dependency idioms, whose corresponding logical structures form the building blocks of Bayesian Network models. The HRA dependency idioms, created in Chapter 4, are summarized below:

- **Causal**: A probabilistic, cause-effect relationship between two HRA constructs (see Section 4.3.1, Figure 4.7).
- **Definitional**: A deterministic relationship wherein the parent directly defines, rather than causes, the child node, and can therefore serve as a distinct state of the child node (see Section 4.3.2, Figure 4.8).
- **Construction**: A probabilistic relationship wherein the parent(s) uncertainly define the child node; this idiom can be used to translate between levels of abstraction (see Section 4.3.3, Figure 4.9).
- **Common Context**: A probabilistic relationship wherein two or more MCFs are dependent by virtue of a common PIF set, i.e., operational context (see Section 4.3.4, Figure 4.10).
- **Situational**: A probabilistic relationship wherein a previous MCF can influence a subsequent MCF by altering the situation (i.e., PIFs) that weigh on the subsequent MCF (see Section 4.3.5, Figure 4.11).
- **Effect-modulating**: A probabilistic relationship that captures the non-linear effects of multiple interacting PIFs (see Section 4.3.6, Figure 4.13).

Research Objective 3: Develop the mathematical framework for quantification of the causal dependency relationships identified in Research Objective 2. RO-3 explicitly addressed the lack of valid and causally-based mathematics for quantifying dependency. The resulting mathematical framework rooted in causality replaces arbitrary dependency mathematics with a causally-informed methodology for computing dependent HEPs.

8.1.3 Key Results: Mathematical Framework (RO-3)

The results of Research Objective 3 provided the methodologies for constructing and quantifying HRA BNs (Chapters 5 and 6, respectively) from HRA data and the HRA dependency idioms.

The principal results of RO-3 are summarized below:

- A traceable methodology for constructing Bayesian networks using qualitative HRA data to build the nodes and the HRA dependency idioms to create the dependency structure, as shown in Figure 5.5 (see Section 5.8).
- An objective, accurate, and data-driven methodology for quantifying the BN using quantitative HRA data, as shown in Figure 6.3 (see Section 6.4).
- A methodology for re-using the idioms and BN fragments from previous models to expedite the construction and quantification of new HRA models, as shown in Figure 6.5 (see Section 6.7).
- A fully-quantified, formative dependency model that demonstrates the above techniques on real HRA data in the nuclear context (see Chapter 6).

Research Objective 3 also created a series of recommendations for improving the field of HRA and the HRA data lifecycle based on the causal, idiom-based framework presented in this research, which are detailed in Chapter 7.

8.2 Research Contributions to Literature

This research has produced several publications that detail the process and results of the research objectives. This section provides a list of the journal and conference publications that resulted from this research, as well as the presentations and workshops that this work facilitated.

8.2.1 Peer-reviewed Journal Publications

This research has produced one peer-reviewed journal publication covering Research Objective 1. A paper discussing Research Objective 2 is currently under second-round review by the journal *Nuclear Science & Engineering*. Papers detailing Research Objective 3 (i.e., Chapters 5 and 6 of this dissertation) is being prepared for publication in a reliability engineering journal. Chapter 7 of this dissertation is being prepared for publication as a perspectives paper.

The following peer-reviewed journal publications have resulted from this dissertation research:

- Vincent P. Paglioni and Katrina M. Groth. "Dependency definitions for quantitative human reliability analysis." In: *Reliability Engineering & System Safety* (2022), p. 108274. DOI: https://doi.org/10.1016/j.ress.2021.108274
- **Vincent P. Paglioni**, and Katrina M. Groth. "Dependency Idioms for Quantitative Human Reliability Analysis." Under second-round review in *Nuclear Science & Engineering*.

The following manuscripts are in preparation for journal submission as a result of this dissertation research:

- **Vincent Philip Paglioni**, and Katrina M. Groth. "Creating Data-Driven HRA BNs with Dependency Idioms, Part I: Construction." In preparation.
- **Vincent Philip Paglioni**, and Katrina M. Groth. "Creating Data-Driven HRA BNs with Dependency Idioms, Part II: Quantification." In preparation.
- **Vincent Philip Paglioni**, and Katrina M. Groth. "Recommendations for HRA Data Lifecycle Management: A Perspective." In preparation.

8.2.2 Refereed Conference Papers and Presentations

In addition to the peer-reviewed journal publications, several refereed conference papers have been published. Refereed conference papers serve as the primary mechanism for engaging engineering community stakeholders. This dissertation has thus far resulted in four full-length, refereed conference papers:

- **Vincent P. Paglioni** and Katrina M. Groth. "Unified Definitions for Dependency in Quantitative Human Reliability Analysis." In: *Proceedings of the 30th European Safety and Reliability Conference and the 15th Probabilistic Safety Assessment and Management Conference*. Venice, Italy, Nov. 1-5, 2020.
- Andres Ruiz-Tagle, **Vincent P. Paglioni**, Enrique Lopez-Droguett, and Katrina M. Groth. "A Framework to Extrapolate and Evaluate Human Reliability Causal Models from Event Report Narratives." In: *2021 International Topical Meeting on Probabilistic Safety Assessment and Analysis (PSA 2021)*. Columbus, OH, Nov. 7-12, 2021.
- **Vincent P. Paglioni** and Katrina M. Groth. "Defining Dependency in HRA." In: *2021 International Topical Meeting on Probabilistic Safety Assessment and Analysis (PSA 2021)*. Columbus, OH, Nov. 7-12, 2021.

- **Vincent Philip Paglioni**, Torrey Mortenson, and Katrina M. Groth. "The human failure event: what is it and what should it be?" In: *Proceedings of the 16th Probabilistic Safety Assessment and Management Conference (PSAM16)*. Honolulu, HI, June 26 - July 1, 2022.

This research has also contributed to two conference papers that are in preparation for the 18th International Probabilistic Safety Assessment and Analysis (PSA 2023) Conference in 2023:

- **Vincent Philip Paglioni** and Katrina M. Groth. "Bridging the Data-Model Gap for HRA: Creating Bayesian Networks from HRA Data" *in preparation for the 18th International Topical Meeting on Probabilistic Safety Assessment and Analysis (PSA 2023)*. Knoxville, TN, July 15-23, 2023.
- **Vincent Philip Paglioni**, Camille S. Levine, Ahmad A. Al-Douri, and Katrina M. Groth. "Why do Human-Machine Teams Fail: Investigating Failure Mechanisms in Human Reliability Analysis" *in preparation for the 18th International Topical Meeting on Probabilistic Safety Assessment and Analysis (PSA 2023)*. Knoxville, TN, July 15-23, 2023.

The conference papers referenced above were presented to technical audiences at their respective conferences. In addition, this dissertation research has been presented at other workshops and venues to further engage with the engineering community:

- **Vincent P. Paglioni** and Katrina M. Groth. "Can HRA Data Address HFE Dependency?" Workshop Presentation. 2020 NRC HRA Data Workshop. Virtual, March 12-13, 2020.
- **Vincent P. Paglioni** and Katrina M. Groth. "Temporal Behaviors of Dependency Relationships in Human Reliability Analysis." Conference Presentation. Society for Risk Analysis Annual Meeting. Virtual, December 13-17, 2020.
- **Vincent Philip Paglioni**, Camille S. Levine, and Katrina M. Groth. "HRA Research – Improving the foundational knowledge of dependency in HRA." Presentation to Sandia National Laboratories. Albuquerque, NM, March 23, 2022.

8.3 Technical Contributions

This dissertation has produced several technical contributions to the field of HRA, and specifically to the art and science of dependency analysis. Each Research Objective has contributed significantly to the HRA body of knowledge, as shown in Figure 8.1. The most impactful technical contributions of this research will be the robust definition of dependency, the comprehensive idiom framework, and the methodology for building and quantifying BNs from HRA data and the HRA dependency idioms.

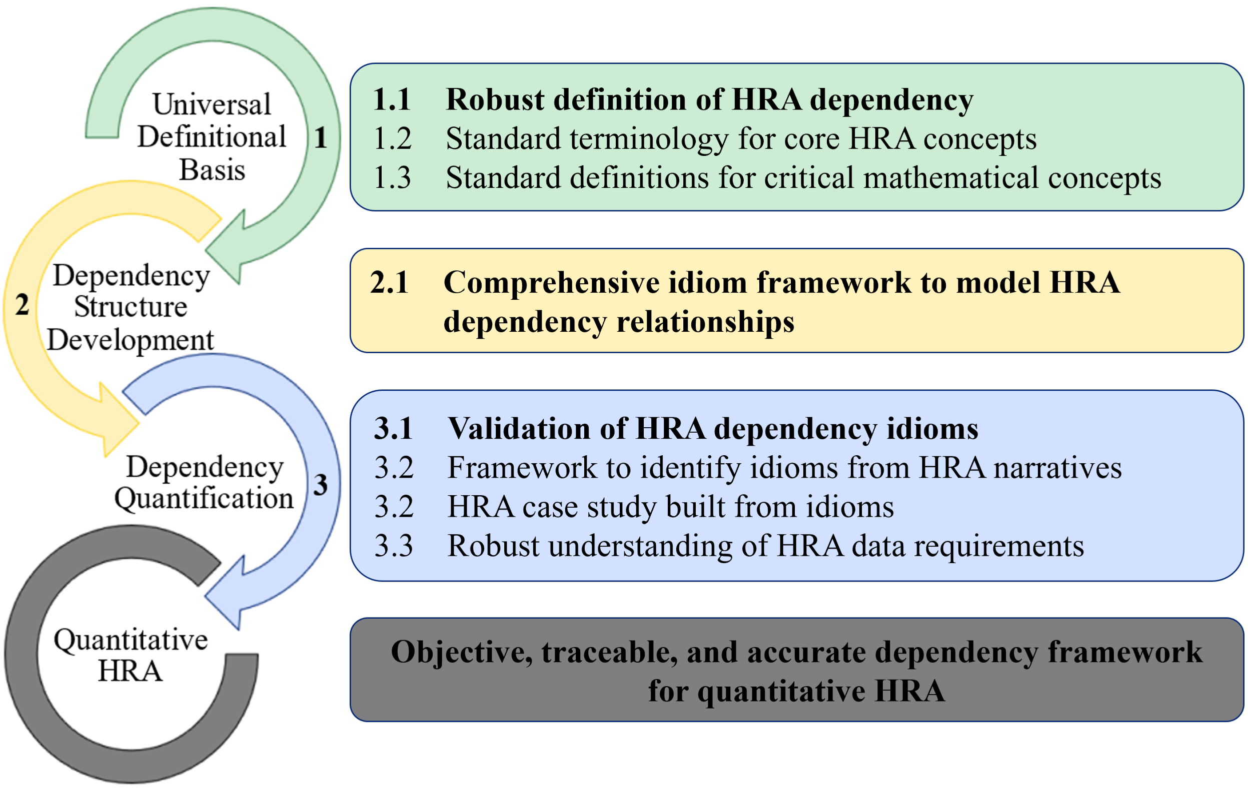


Figure 8.1: Technical contributions imparted by each Research Objective.

8.4 Future Work

Future work in this area will be undertaken across three main thrusts, briefly discussed below. These areas of critical exploration will utilize the principal findings of this research – the definitions, idioms, construction and quantification methodologies, and recommendations – to continue improving the theoretical foundations and developing practical applications for HRA.

1. Building comprehensive HRA causal models with the definitions and idioms
2. Applying the idioms to novel contexts for HRA, including external hazards
3. Designing a robust data collection framework informed by this research
4. Developing the methodology to use the definitions and idioms within existing HRA models

8.4.1 Build comprehensive HRA causal models

The definitions, HRA dependency idioms, and methodologies created in this dissertation should be used to build the next generation of robust HRA models for multiple applications. This dissertation enables the creation of objective, traceable, and accurate causal models that can be used for diagnostic and prognostic purposes, and ultimately provide a more comprehensive understanding of why, how, and how often human error occurs. While the definitions and HRA dependency idioms on their own can improve the understanding of human error in complex engineering systems, building robust causal HRA models is necessary to gain actionable insights for improving human reliability. Therefore, it is essential that the results of this dissertation be used to build and quantify full models for HRA that capture the full, causal picture of HMT performance and reliability.

Building comprehensive HRA models from the results of this dissertation will have many benefits, including increasing the traceability and objectivity of HRA . Building models from the constructs defined in this dissertation will ensure that the model captures the aspects of a scenario that must be included in a model. Furthermore, this ensures that the models are broadly understandable and amenable to using both expert elicitation and HRA data sources, especially if they are collected under the recommendations provided in Chapter 7. The definitions provided in this dissertation can also guide analysts in defining unique constructs observed in the model application; for instance, if new MCFs or CFMs are required to understand HRA for external hazards [121]. The definitions provided in this dissertation will thus ensure that the next generation of HRA models is cohesive, traceable, and sound.

The HRA dependency idioms created in this dissertation will similarly facilitate building the next generation of HRA models, ones based on the causal, probabilistic structure engendered through BNs. The idioms provide the logical framework for assessing dependency relationships in addition to the discrete graphical structures that form the building blocks of the ultimate BN. The HRA dependency idioms allow analysts to assemble a robust model from a collection of modeling constructs by putting dependency – the causal, probabilistic relationships that govern reality – at the forefront of the model-making process. This is a significant improvement over current HRA modeling techniques, which rely on checklists and other less-traceable model constructions. Because the HRA dependency idioms are defined as general causal structures between generalized HRA constructs, rather than specific PIFs or MCFs, they can be used to construct models of both existing and novel contexts. The HRA dependency idioms therefore facilitate the traceable creation of comprehensive HRA models from the standard set of constructs defined in this dissertation, and set up a model that can be quantified objectively and accurately using HRA data for any context.

The models created from the definitions and HRA dependency idioms can model both current and novel contexts to glean a more comprehensive understanding of human reliability. However, maximizing the utility of these models requires quantifying them to understand the impacts of dependency. This dissertation created a data-driven, objective process for parameterizing model CPTs from data, and thereby quantifying the HRA models. The methodology developed in this dissertation requires no expert elicitation in the quantification phase, even when the available data is extremely sparse. This is a significant improvement over the current generation of HRA methods, which all generally require expert elicitation at some phase of quantification, and rely on unfounded and spurious adjustments to probabilities to reflect the effects of dependency. Instead, this methodology uses the idiom-based formative dependency structure to guide model quantification, which is rooted in available data.

The definitions created in this dissertation provide a common, uniform technical basis for producing comprehensive HRA models that can be easily traced, validated, and understood. The HRA dependency idioms provide a coherent, logical structure to HRA dependency and the ability to create formative dependency models in a traceable manner. Finally, the construction and quantification methodologies allow for the traceable and objective construction of full-scenario models and accurate quantification based on HRA data. The HRA models built from this dissertation can replace the existing generation(s) of HRA models to glean new insights for HRA in existing applications and extend to new applications to understand human reliability in novel domains.

8.4.2 Apply the HRA Dependency Idioms to Novel Contexts

The scope and scale of complex engineering systems are evolving rapidly with the advent of advanced technologies including automation, artificial intelligence, and machine learning. Additionally, novel contexts for human-operated systems are being explored that include long-term spaceflight and extra-planetary habitation. Finally, natural hazards are posing an increasing threat to the viability of terrestrial critical infrastructure systems as their severity, location, and predictability are being impacted by climate change. All of these factors may necessitate changes in the role, capabilities, and requirements for human operators, which must be accounted for in HRA. The principal results of this research, namely the definitions and idioms, provide a robust basis to investigate these new areas for HRA.

External hazards, particularly natural hazards (e.g., hurricanes, flooding, and other stochastic external events), represent challenging and unique scenarios for understanding human reliability in complex engineering systems. Such events may require ex-control room actions, which are not typically accounted for in HRA and which require the inclusion of additional PIFs, CFMs, and MCFs in order to model. As a result, existing HRA methods are not always applicable to such scenarios. Thus, there is a significant need to develop HRA methods capable of handling the dynamic environments of natural hazards, action-based and cognition-based human functions, and potentially novel constructs imposed by these scenarios. The definitions provided in Chapter 3 can serve as a starting point for defining any new constructs encountered during the study of natural hazards (or ex-control room) HRA. Further, the HRA dependency idioms (Chapter 4) will provide the causal structure of dependencies even in ex-control room HRA. The definitions, HRA dependency idioms, and construction and quantification methodologies developed in this dissertation should be used to traceably and soundly create HRA causal models of, e.g., external hazards at nuclear power plants.

The development and deployment of operator-assistive technologies (e.g., AI/ML-enabled diagnostics and prognostics, automation, and remote operation) could represent a significant shift in the role of the operator for complex engineering systems, but are unlikely to invalidate the use of human personnel completely. As a result, HRA must adapt pre-emptively as part of robust a risk analysis to ensure the reliability and resiliency of humans in these new roles prior to their implementation in the field. This will require investigating the application of existing idioms for these novel roles and using them to collect data regarding these adapted roles. It is expected that the idioms will still be applicable to changing operator roles - again, the idioms represent the reality of causality rather than construct-specific relationships. However, the advent of new operational roles will likely explicate novel constructs (PIFs, CFMs, and MCFs) that must be investigated in the context of the idiom framework. The idioms can be used to bridge HRA theory and application, allowing researchers to posit the new constructs and causal relationships and identify/verify the theory with data collected under an idiom-informed protocol. Further, the idioms provide pre-built causal structures that can facilitate rapid model creation and alteration. Therefore, using the idioms allows analysts to understand the HRA implications of novel roles and interface system designs during even the early stages of system development.

The exploration of space is rapidly advancing toward the possibilities of commercialized spaceflight, long-term crewed space missions (e.g., to Mars) and extra-planetary surface habitation (e.g., on the Moon and/or Mars). Although HRA has been investigated for existing space operational paradigms (e.g., short-term orbital missions, shuttle operation), there are significant differences in context that will require an investigation of both the theoretical basis and practical application of HRA. Spaceflight already represents a significantly more complex context than terrestrial system operation, due to severe data limitations and a poor understanding of the context. The PIFs, their potentially (likely) unique impact to human performance, and the CFMs associated with spaceflight are not well understood, and adjusting from the existing spaceflight context to new types of missions may exacerbate these issues further. As with operator-assistive technologies, the successful development and deployment of new spaceflight systems will require a thorough understanding of the associated human reliability, which can be investigated using the definitions and idioms from this research as useful starting points. Theorizing, identifying, and validating new PIFs, CFMs, and

MCFs should use the definitions provided in Chapter 3 as a basis and guidance. Analysts can use the idioms as a guide to investigating the causal structure that may link MCFs in spaceflight, and use the idioms to guide the collection of data from spaceflight simulators, interviews, event reports, and other historical sources to begin building an understanding of HRA in novel contexts.

8.4.3 Design the Improved Data Collection Framework

Data is critical to ensure that the HRA field has a robust understanding of, and modeling methods for, human-machine team reliability within complex systems. However, as this research has indicated, currently-available HRA data is not necessarily amenable to furthering either of these focuses of HRA. The use of inconsistent definitions of HRA constructs in data and methods complicates processes for learning from data and using it in modeling applications. Chapter 7 outlined ways that this research can be used to improve HRA and the data lifecycle.

Making recommendations for improving HRA data across its lifecycle only sets guidelines. Future research will be required to design an improved data collection framework that will support future modeling and quantification efforts, based on the definitions and idioms in this dissertation. This dissertation should inform the constructs that should be collected, how data is stored to maximize the benefits to modeling efforts. However, the methodology for actually collecting the data – including how to collect specific constructs, the sensors/collection methodologies that should be used – will be the subject of future research.

8.4.4 Develop Methodology to Use Definitions and Idioms in Existing HRA

Incorporating the definitions and idioms from this research into existing HRA methods can, as explained in Chapter 7, improve the traceability, objectivity, and veracity of HRA methods in addition to providing formative dependency modeling capabilities. Incorporating the definitions into existing HRA can be done by revising existing methods to align their construct and concept definitions more closely with those provided here. This would strengthen the theoretical and technical basis of existing HRA. Moreover, this would facilitate additional cross-method research and communication to improve the comprehensive understanding of HRA. Allowing disparate methods to be discussed with a common basis will help to reduce the inter-method variability that is engendered by inconsistent definitions and usage of HRA constructs.

Using the HRA dependency idioms in existing HRA methods can improve the traceability and objectivity of dependency analysis, by explicating the models and enforcing causality-based dependency relationships over correlational connections. The idioms may be incorporated almost immediately into BN-based HRA methods such as Phoenix and IDHEAS. In such methods, using the idioms would provide additional structure and traceability to the BN construction process and imbue the models with formative dependency. However, the idioms may not be directly usable in every extant HRA method, particularly if the method is to be used without significant modification. Checklist-based (i.e., non-model based) HRA methods like SPAR-H and IDHEAS have been represented with BNs that copy the existing quantification structures [63, 65], but do not change the dependency modeling or quantification. One of the main benefits of the idioms is to empower the dependency quantification through data-driven processes. However, the idioms could be used qualitatively as a mechanism for tracing, validating, and improving the dependency relationships included in non-model based HRA methods.

Future work in this area should investigate techniques to incorporate idiom structures into existing HRA models, in both verification/validation/improvement of existing dependency techniques and in replacing or augmenting BN representations of HRA methods. For BN-based methodologies, case studies should be performed that show the efficacy of the idiom-based BN for replacing and/or augmenting the existing BN structure. For non-model based HRA methods, the idioms should be used to trace and validate existing dependency relationships and quantification techniques. Further work should investigate how the idioms can be integrated with existing dependency processes in non-model based HRA methods, to form either hybrid BN and checklist methods or fully replacing the checklist dependency methods with idiom-based BNs.

8.5 Broader Research Impact

This research follows in a long tradition of seeking explanatory theories for "human error" and significantly expands our understanding of why and how humans fail by recognizing the complex causal relationships that connect human-machine team performance to the system context. In that sense, this dissertation represents the latest step in a chain of philosophical inquiry that stretches at least 2,000 years to Seneca the Younger.

In another sense, the research contained herein radically changes the perception and modeling of human error by enforcing causality as the central explanatory theory of dependency and considering the human-machine team as a singular organism. The causal conceptualization of dependency developed in this research roots explanations of human error in real, physical relationships and provides a scientific, analytical basis to the modeling of human reliability. This revolutionary shift from first-order theories based in coincidence to rational, causal models informed by literature and data is a step towards a robust, realistic understanding of human-machine teams in operational contexts. As technology continues to advance and evolve, the human role in complex engineering systems will continue to change, and it is imperative to gain an understanding of the human contribution to risk in these systems. The causal conceptualization of dependency established in this research is critical to building this understanding and ultimately to effectively reducing human-induced failures in complex engineering systems.

This research provided standard definitions for many terms and concepts critical to the study and performance of HRA, and created a robust definition for dependency itself that is rooted in causality and probability. These results standardize the foundations of the field and will allow easier cross-communication between future researchers using these construct definitions, even in different HRA methods. This will help to reduce the inconsistencies in HRA methods that drive quantitative variations in HRA results, and thus improve our understanding of system and human reliability. Building on this success, this research then created the dependency idioms that exist between HRA constructs, that provide structure to the assessment of dependency. The idioms, and their associated BN structures, ensure that HRA dependency assessments will be guided by a logical structure, result in traceable causal networks, and be quantified in an objective manner based on data. This research also developed a structured methodology for creating data-driven BN models that enforce formative dependency and improve the accuracy, traceability, and objectivity of HRA. Finally, the recommendations offered by this research for improving HRA will ensure a robust data supply and a more coherent field.

Humans, by virtue of their humanity, will continue to be both fallible beings and critical components of complex engineering systems that hold key roles in design, operation, and maintenance, likely in increasingly symbiotic roles with the system itself. Understanding how human-machine teams fail in their duties, how environmental, system, and personal factors can affect the proclivity towards failure, and how design and operational principles can be improved to support reliability, are critical to ensuring the continued safe and reliable operation of critical systems. This research provides the basis for this understanding by developing the theoretical causal framework within which human reliability should be viewed and analyzed. Using the idiom structures from this research will allow analysts to verify and validate existing HRAs, build improved causal models to support future PRAs, and provide the framework to effectively collect data to support these assessments.

In summary, this research helps solve a central problem within quantitative HRA, laid out by James Reason in 1990 [105, p. 233]:

Without [...] a workable theoretical infrastructure, there can be little or no principled basis to the business of human reliability quantification.

Appendix A: Sample SACADA Dataset

The purpose of this appendix is to provide an example of the SACADA data which was used in the performance of the dissertation research. An overview of the SACADA framework is provided in [84].

Tables A.1 – A.7 show the SACADA data recorded for crews in the simulation of the ISLOCA scenario used for the case study in Chapters 5 and 6. These tables show the PIFs recorded for all TOEs regardless of outcome parity (e.g., SAT, UNSAT, etc.), a subset of which were retained in the BN model built in this dissertation.

Table A.1: The first and last columns of SACADA data, showing the name information for each TOE in the ISLOCA scenario, along with the outcomes and UNSAT Ratio.

Scenario	Year	Cycle	Malf. Order	TOE Order	UNSAT	SAT	SAT Δ	SAT+	Total	UNSAT Ratio(%)
—	—	—	—	—	UNSAT	SAT	SAT Δ	SAT+	Total	UNSAT Ratio(%)
ISLOCA	2014	1	1	1	0	1	1	0	2	0
ISLOCA	2014	1	1	2	0	1	1	0	2	0
ISLOCA	2014	1	1	3	1	1	0	0	2	50
ISLOCA	2014	1	1	4	0	2	0	0	2	0
ISLOCA	2014	1	1	5	0	2	0	0	2	0
ISLOCA	2014	1	1	6	1	1	0	0	2	50
ISLOCA	2014	1	1	7	0	2	0	0	2	0
ISLOCA	2014	1	1	9	0	1	0	0	1	0
ISLOCA	2014	1	1	10	0	2	0	0	2	0
ISLOCA	2014	1	1	11	1	1	0	0	2	50
ISLOCA	2014	1	1	11	0	2	0	0	2	0
ISLOCA	2014	1	1	12	0	1	1	0	2	0
ISLOCA	2014	1	1	15	0	1	0	0	1	0

Table A.2: Columns H – J of SACADA data, showing the MCF name, importance, and cognitive type.

TOE Order	TOE (training objective element)	Importance	Cognitive Type
—		0:NULL 1:Other 2:Significant 3:Safety Significant 4:Critical	0:NULL 1:Monitoring/Detection 2:Diagnosis & Response Planning 3:Manipulation 4:External Communication
1	Decision/Diagnosis	2	2
2	Action Taking	2	3
3	Decision/Diagnosis	2	2
4	Decision/Diagnosis	3	2
5	Communication	3	4
6	Information Gathering	3	1
7	Decision/Diagnosis	3	2
9	Action Taking	3	0
10	Action Taking	3	3
11	Decision/Diagnosis	3	2
11	Decision/Diagnosis	3	2
12	Action Taking	3	3
15	Decision/Diagnosis	3	2

Table A.3: Columns K – Q of SACADA data, showing the last I-phase relevant PIFs.

TOE Order	Monitoring/ Detection *Detection Type*	Alarms/Status Tile *Detection Mode*	Alarms/ Status Tile *Status of Alarm* Board	Alarms/Status Tile *Expectation of Alarm/Indication Change*	Meter/Light/Flag *Detection Mode*	M/L/F *Individual Indicator*	M/L/F *Mimics/ Display* etc.
—	0:NULL 1:Alarm 2:Status Tile 3:Meter 4:Indication Light 5:Flag 6:Computer 7:Other	0:NULL 1:Self-Revealing 2:Procedure Directed Check 3:Procedure Directed Monitoring 4:Awareness/ Inspection	0:NULL 1:Dark 2:Busy 3:Overloaded	0:NULL 1:Expected 2:Not Expected 3:Not Applicable	0:NULL 1:Procedure Directed Check 2:Knowledge-Driven Monitoring 3:Procedure-Directed Monitoring 4:Awareness/ Inspection	0:NULL 1:Slight Change 2:Distinct Change	0:NULL 1:No Mimics 2:Small Indications 3:Similar Displays
1	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0
6	1	1	2	2	0	0	0
7	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0
11	0	0	0	0	0	0	0
11	0	0	0	0	0	0	0
12	0	0	0	0	0	0	0
15	0	0	0	0	0	0	0

Table A.4: Columns R – W of SACADA data, showing the first D-phase relevant PIFs.

TOE Order	Diagnosis and Response Planning *Diagnosis or Response Planning*	Diagnosis *Diagnosis Basis*	Diagnosis *Familiarity*	Diagnosis *Outcome*	Diagnosis *Information Integration*	Diagnosis *Information Specificity*
—	0:NULL 1:Primarily Diagnosis 2:Primarily Response Planning/ Decision Making	0:NULL 1:Procedure 2:Skill 3:Knowledge	0:NULL 1:Standard 2:Novel 3:Anomaly	0:NULL 1:Procedure-Based Activity 2:Skill-Based Behavior 3:Knowledge-Based Behavior	0:NULL 1:Timing of Information 2:Ambiguous Information 3:Integration Required	0:NULL 1:Specific 2:Not Specific 3:Not Applicable
1	1	2	0	0	0	0
2	0	0	0	0	0	0
3	2	0	0	0	0	0
4	2	0	0	0	0	0
5	0	0	0	0	0	0
6	0	0	0	0	0	0
7	2	0	0	0	0	0
9	0	0	0	0	0	0
10	0	0	0	0	0	0
11	2	0	0	0	0	0
11	2	0	0	0	0	0
12	0	0	0	0	0	0
15	2	0	0	0	0	0

Table A.5: Columns X – AB of SACADA data, showing the last D-phase relevant PIFs.

TOE Order	Diagnosis Information Quality	Response Planning/ Decision Making *Decision Basis*	Response Planning/ Decision Making *Familiarity*	Response Planning/ Decision Making *Uncertainty*	Response Planning/ Decision Making *Outcome*
—	0:NULL 1:Missing Information 2:Misleading Information 3:Conflicting Information	0:NULL 1:Procedure 2:Skill 3:Knowledge	0:NULL 1:Standard 2:Adaptation Required 3:Anomaly	0:NULL 1:Clear 2:Uncertain 3:Competing Priorities 4:Conflicting Guidance	0:NULL 1:Procedure-Based Activity 2:Skill-Based Behavior 3:Knowledge-Based Behavior
1	0	0	0	0	0
2	0	0	0	0	0
3	0	1	1	1	0
4	0	3	2	3	0
5	0	0	0	0	0
6	0	0	0	0	0
7	0	1	2	4	0
9	0	0	0	0	0
10	0	0	0	0	0
11	0	3	3	1	0
11	0	1	1	2	1
12	0	0	0	0	0
15	0	1	1	2	0

Table A.6: Columns AC – AG of SACADA data, showing the last D-phase relevant PIFs.

TOE Order	Manipulation *Type of Action*	Manipulation *Location*	Manipulation *Guidance*	Manipulation *Recoverability*	Manipulation *Additional Factors*
—	0:NULL 1:Simple and Distinct 2:Order 3:Maintaining	0:NULL 1:Main or Auxiliary Control Board 2:Back Control Panels	0:NULL 1:Procedure 2:Skill of the Craft (Non-Faulted Hardware) 3:STAR (Faulted Hardware)	0:NULL 1:Immediately Recoverable 2:Recoverable With Significant Efforts 3:Unrecoverable	0:NULL 1:Unintuitive Plant Response 2:Unintuitive Controls 3:Additional Mental Effort Required 4:Inadequate Feedback 5:Similar Controls
1	0	0	0	0	0
2	3	1	2	3	0
3	0	0	0	0	0
4	0	0	0	0	0
5	0	0	0	0	0
6	0	0	0	0	0
7	0	0	0	0	0
9	0	0	0	0	0
10	2	1	2	1	0
11	0	0	0	0	0
11	0	0	0	0	0
12	3	1	1	1	5
15	0	0	0	0	0

Table A.7: Columns AH – AM of SACADA data, showing the C-phase relevant and Overarching PIFs.

TOE Order	Communication Between Crew & Simulator Booth *Communication Driver*	Communication Between Crew & Simulator Booth *Direction of Communication*	Overarching Issues *Workload*	Overarching Issues *Time Criticality*	Overarching Issues *Extent of Comms Required*	Overarching Issues *Other Demands/ Factors*
—	0:NULL 1:Specifically Procedure Directed 2:Not Specifically Driven	0:NULL 1:From Booth 2:To Booth 3:Public Address Announcement 4:Other	0:NULL 1:Normal 2:Concurrent Demands 3:Multiple Concurrent Demands	0:NULL 1:Expansive Time Available 2:Nominal Time Available 3:Barely Adequate Time Available	0:NULL 1:Nominal Comm. 2:Extensive Onsite Comm 3:Extensive Comm Within Control Room	0:NULL 1:Non-Standard 2:Noisy Bkgd 3:Coordination 4:Communicator Unavailable 5:Multiple Demands 6:Memory Demands
1	0	0	1	2	1	0
2	0	0	1	2	1	0
3	0	0	1	2	1	0
4	0	0	1	3	3	0
5	2	2	2	2	1	1
6	0	0	2	2	1	1
7	0	0	2	2	1	1
9	0	0	2	2	1	1
10	0	0	2	2	1	1
11	0	0	1	2	1	1
11	0	0	2	2	1	1
12	0	0	2	2	1	5
15	0	0	2	2	2	5

Appendix B: Selected Conditional Probability Tables

This appendix provides CPTs for nodes in the I, A, and C cognitive phases, as well as the overarching nodes in the model.

B.1 I-phase Submodel CPTs

Tables B.1 – B.4 provide CPTs for the PIFs and MCF node in the I-phase submodel.

Table B.1: CPT for the PIF Meter/Light/Flag Individual Indicator in the I-phase submodel.

PIF State	**Probability**
Slight Change	0.4544
Distinct Change	0.5456

Table B.2: CPT for the PIF Meter/Light/Flag Detection Mode in the I-phase submodel.

PIF State	**Conditional Probability**	
Meter/Light/Flag Individual Indicator	*Slight Change*	*Distinct Change*
Procedure Directed Check	0.4847	0.4291
Knowledge Driven Monitoring	0.1636	0.1821
Procedure Directed Monitoring	5.5021e-07	0.1072
Awareness Inspection	0.3517	0.2816

Table B.3: CPT for the PIF Overarching Issues Time Criticality in the I-phase submodel.

PIF State	**Conditional Probability**					
Overarching Issues Workload	*Normal*		*Concurrent Demands*		*Multiple Concurrent Demands*	
Previous MCF	*SAT*	*UNSAT*	*SAT*	*UNSAT*	*SAT*	*UNSAT*
Expansive Time	0.6805	0.6659	0.0790	0.0521	0.0877	8.6266e-05
Nominal Time	0.2420	0.2095	0.8402	0.8975	0.6823	0.3513
Barely Adequate Time	0.0775	0.1245	0.0808	0.0505	0.2299	0.6486

Table B.4: Partial CPT for the I-phase MCF. Only the first 18 cells are shown here because there are 144 cells in the full CPT.

MCF State	**Conditional Probability**									
Meter/Light/Flag Detection Mode	*Procedure Directed Check*									→
Meter/Light/Flag Ind. Indicator	*Slight Change*									→
Overarching Issues Workload	*Normal*			*Concurrent Demands*			*Multiple Concurrent Demands*			→
Overarching Issues Time Criticality	*Expansive*	*Nominal*	*Barely Adequate*	*Exp.*	*Nom.*	*B.A*	*Exp.*	*Nom.*	*B.A*	→
SAT	0.9933	0.9999	0.9999	0.9999	0.9905	0.9999	0.9998	0.9999	0.9999	→
UNSAT	0.0067	3.5794e-05	0.0001	4.9363e-05	0.0095	4.8305e-05	0.0002	2.2904e-05	6.4868e-05	→

B.2 A-phase Submodel CPTs

Tables B.5 – B.12 provide CPTs for the PIFs and MCF in the A-phase submodel.

Table B.5: CPT for the PIF Manipulation Guidance in the A-phase submodel.

PIF State	**Probability**
Procedure	0.9499
Skill of the Craft Non-faulted Hardware	0.0279
STAR Faulted Hardware	0.0222

Table B.6: CPT for the PIF Manipulation Type of Action in the A-phase submodel.

PIF State	**Probability**
Simple & Distinct	0.4149
Order	0.4301
Maintaining	0.1550

Table B.7: CPT for the PIF Manipulation Location in the A-phase submodel.

PIF State	**Probability**
Main or Auxiliary Control Board	0.9741
Back Control Panels	0.2587

Table B.8: CPT for the PIF Manipulation Recoverability in the A-phase submodel.

PIF State	**Conditional Probability**		
Manipulation Guidance	*Procedure*	*Skill of the Craft Non-faulted Hardware*	*STAR Faulted Hardware*
Immediately Recoverable	0.4194	0.3592	0.4029
Recoverable with Significant Effort	0.39481	0.4709	0.1902
Unrecoverable	0.1858	0.1699	0.4069

Table B.9: Partial CPT for the PIF Overarching Issues Time Criticality in the A-phase submodel, when the type of action is "simple and distinct."

PIF State	**Conditional Probability**					
Manipulation Type of Action	*Simple and Distinct*					
Overarching Issues Workload	*Normal*		*Concurrent Demands*		*Multiple Concurrent Demands*	
Previous MCF	*SAT*	*UNSAT*	*SAT*	*UNSAT*	*SAT*	*UNSAT*
Expansive Time	0.5741	0.3400	0.1565	0.0622	5.5504e-06	0.0003
Nominal Time	0.4258	0.6598	0.7812	0.8223	0.4946	0.6179
Barely Adequate Time	3.1076e-06	0.0001	0.0622	0.1156	0.5054	0.3818

Table B.10: Partial CPT for the PIF Overarching Issues Time Criticality in the A-phase submodel, when the type of action is "order."

PIF State	**Conditional Probability**					
Manipulation Type of Action	*Order*					
Overarching Issues Workload	*Normal*		*Concurrent Demands*		*Multiple Concurrent Demands*	
Previous MCF	*SAT*	*UNSAT*	*SAT*	*UNSAT*	*SAT*	*UNSAT*
Expansive Time	0.6078	0.1982	0.1534	0.0789	5.4432e-06	0.0003
Nominal Time	0.3259	0.2997	0.8071	0.6897	0.5662	0.4899
Barely Adequate Time	0.0663	0.5021	0.0395	0.2314	0.4338	0.5099

Table B.11: Partial CPT for the PIF Overarching Issues Time Criticality in the A-phase submodel, when the type of action is "maintaining."

PIF State	**Conditional Probability**					
Manipulation Type of Action	*Maintaining*					
Overarching Issues Workload	*Normal*		*Concurrent Demands*		*Multiple Concurrent Demands*	
Previous MCF	*SAT*	*UNSAT*	*SAT*	*UNSAT*	*SAT*	*UNSAT*
Expansive Time	0.8023	0.9992	0.0925	0.0002	1.5006e-05	0.0007
Nominal Time	0.1977	0.0004	0.8226	0.9996	0.9999	0.9986
Barely Adequate Time	8.4572e-06	0.0004	0.0849	0.0002	1.5006e-05	0.0007

Table B.12: Partial CPT for the A-phase MCF. Only the first 18 cells are shown here because there are 968 cells in the full CPT.

MCF State	**Conditional Probability**									
Manipulation Location	*Main or Auxiliary Control Boards*									→
Manipulation Guidance	*Procedure*									→
Manipulation Recoverability	*Immediately Recoverable*									→
Overarching Issues Workload	*Normal*			*Concurrent Demands*			*Multiple Concurrent Demands*			→
Overarching Issues Time Criticality	*Expansive*	*Nominal*	*Barely Adequate*	*Exp.*	*Nom.*	*B.A*	*Exp.*	*Nom.*	*B.A*	→
SAT	0.9999	0.9999	0.5	0.9625	0.9875	0.9999	0.5	0.9748	0.9999	→
UNSAT	2.063e-05	2.725e-05	0.5	0.0375	0.0125	8.187e-05	0.5	0.0252	4.235e-05	→

B.3 C-phase Submodel CPTs

Tables B.13 – B.16 provide CPTs for the PIFs and MCF in the C-phase submodel.

Table B.13: CPT for the PIF Communication between Crew and Simulator Booth Driver in the C-phase submodel.

PIF State	Probability
Specifically Procedure Directed	0.29455
Not Specifically Driven	0.7055

Table B.14: CPT for the PIF Communication between Crew and Simulator Booth Direction in the C-phase submodel.

PIF State	**Conditional Probability**	
Communication Between Crew and Simulator Booth Driver	*Specifically Procedure Driven*	*Not Specifically Driven*
From Booth	8.4891e-07	0.0269
To Booth	0.6658	0.8194
Public Address Announcement	0.2805	0.0589
Other	0.0536	0.0948

Table B.15: CPT for the PIF "Overarching Issues Time Criticality" in the C-phase submodel.

PIF State	**Conditional Probability**					
Overarching Issues Workload	*Normal*		*Concurrent Demands*		*Multiple Concurrent Demands*	
Previous MCF	*SAT*	*UNSAT*	*SAT*	*UNSAT*	*SAT*	*UNSAT*
Expansive Time	0.9476	0.9999	0.1826	0.0584	2.3224e-06	0.0001
Nominal Time	0.0064	6.4826e-05	0.6885	0.7590	0.3761	0.1997
Barely Adequate Time	0.0460	6.4826e-05	0.1289	0.1826	0.6239	0.8002

Table B.16 does not provide the "first" 18 cells in the table, as is shown in the other MCF CPTs. The additional sparsity relating to C-phase MCFs produced a CPT where many probabilities were equal to 0.5, indicating an absence of data relating to specific combinations of the C-phase PIFs. A portion of the CPT with more representative values has been provided.

Table B.16: Partial CPT for the C-phase MCF. Only a selection of 18 cells are shown here because there are 144 cells in the full CPT.

MCF State	**Conditional Probability**									
Communication Driver	*Specifically Procedure Driven*									$\rightarrow$
Communication Direction	*To Booth*									$\rightarrow$
Overarching Issues Workload	*Normal*			*Concurrent Demands*			*Multiple Concurrent Demands*			$\rightarrow$
Overarching Issues Time Criticality	*Expansive*	*Nominal*	*Barely Adequate*	*Exp.*	*Nom.*	*B.A*	*Exp.*	*Nom.*	*B.A*	$\rightarrow$
SAT	0.9999	0.9985	0.9998	0.9999	0.9999	0.9999	0.5	0.9999	0.9799	$\rightarrow$
UNSAT	1.0277e-05	0.0015	0.0002	2.3858e-05	6.2730e-06	3.2819e-05	0.5	4.6559e-05	0.0201	$\rightarrow$

B.4 Overarching CPTs

Tables B.17 – B.22 provide CPTs for the PIF Overarching Issues Workload, the cognitive phase CSP nodes, and the HFE node.

Table B.17: CPT for the overarching PIF Overarching Issues Workload. This PIF was assumed to be a static aspect of the scenario: it does not change over time and has no parent nodes.

PIF State	**Probability**
Normal	0.2617
Concurrent Demands	0.5918
Multiple Concurrent Demands	0.1465

Table B.18: CPT for the I-phase cognitive sub-phase (CSP) node, which is definitionally dependent on the I-phase MCF node.

CSP State	**Conditional Probability**	
I-phase MCF (6)	*SAT*	*UNSAT*
No HFE	1	0
HFE	0	1

Table B.19: Partial CPT for the D-phase CSP node, which is definitionally dependent on the D-phase MCF nodes. As a result, the CSP node only takes state ”No HFE” if all D-phase MCFs are SAT. If any D-phase MCF is failed, the CSP registers an HFE.

CSP State	**Conditional Probability**								
D-phase MCF (1)	*SAT*								→
D-phase MCF (3)	*SAT*								→
D-phase MCF (4)	*SAT*								→
D-phase MCF (7)	*SAT*								→
D-phase MCF (10)	*SAT*				*UNSAT*				→
D-phase MCF (11)	*SAT*		*UNSAT*		*SAT*		*UNSAT*		→
D-phase MCF (13)	*SAT*	*UNSAT*	*SAT*	*UNSAT*	*SAT*	*UNSAT*	*SAT*	*UNSAT*	→
No HFE	1	0	0	0	0	0	0	0	→
HFE	0	1	1	1	1	1	1	1	→

Table B.20: Partial CPT for the A-phase CSP node, which is definitionally dependent on the A-phase MCF nodes. As a result, the CSP node only takes state ”No HFE” if all A-phase MCFs are SAT. If any A-phase MCF is failed, the CSP registers an HFE.

CSP State	**Conditional Probability**								
A-phase MCF (2)	*SAT*								$\rightarrow$
A-phase MCF (8)	*SAT*				*UNSAT*				$\rightarrow$
A-phase MCF (9)	*SAT*		*UNSAT*		*SAT*		*UNSAT*		$\rightarrow$
A-phase MCF (12)	*SAT*	*UNSAT*	*SAT*	*UNSAT*	*SAT*	*UNSAT*	*SAT*	*UNSAT*	$\rightarrow$
No HFE	1	0	0	0	0	0	0	0	$\rightarrow$
HFE	0	1	1	1	1	1	1	1	$\rightarrow$

Table B.21: CPT for the C-phase cognitive sub-phase (CSP) node, which is definitionally dependent on the C-phase MCF (MCF 5).

CSP State	**Conditional Probability**	
C-phase MCF (5)	*SAT*	*UNSAT*
No HFE	1	0
HFE	0	1

Table B.22: CPT for the Objective node, which is definitionally dependent on the CSP nodes.

Objective State	**Conditional Probability**															
I CSP	*No HFE*								*HFE*							
D CSP	*No HFE*				*HFE*				*No HFE*				*HFE*			
A CSP	*No HFE*		*HFE*		*No HFE*		*HFE*		*No HFE*		*HFE*		*No HFE*		*HFE*	
C CSP	*No HFE*	*HFE*	*No HFE*	*HFE*	*No HFE*	*HFE*	*No HFE*	*HFE*	*No HFE*	*HFE*	*No HFE*	*HFE*	*No HFE*	*HFE*	*No HFE*	*HFE*
No HFE	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
HFE	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

Bibliography

[1] IAEA (International Atomic Energy Agency). *Nuclear Power Plant Operating Experience from the IAEA/NEA International Reporting System for Operating Experience 2015-2017*. Tech. rep. Vienna: International Atomic Energy Agency, 2020, pp. 1–68.

[2] April M. Whaley, Jing Xing, Ronald L. Boring, Stacy M. L. Henrickson, Jeffrey C. Joe, Katya L. Le Blanc, and Stephanie L. Morrow. *Cognitive Basis for Human Reliability Analysis (NUREG - 2114)*. Tech. rep. U.S. Nuclear Regulatory Commission, 2016.

[3] NuScale Power, LLC. *NuScale Standard Plant Design Certification Application Chapter 18: Human factors Engineering*. Tech. rep. 2020.

[4] Oklo, Inc. *Oklo Power Combined Operating License Application for the Aurora at INL Part II : Final Safety Analysis Report*. Tech. rep. Sunnyvale, 2020.

[5] TerraPower, LLC. *TerraPower's Traveling Wave Technology: Bringing Advanced nuclear to Market*. Bellevue, 2021.

[6] Piyush Garg, Yunfei Zhao, and Carol Smidts. "Human Performance Shaping Factors for Micro Nuclear Reactors." In: *Nuclear Plant Instrumentation, Control, and Human-Machine Interface Technologies (NPIC&HMIT 2019)* (Feb. 9–14, 2019). Orlando, FL, 2019, pp. 1195–1209.

[7] Vincent P. Paglioni and Katrina M. Groth. "Defining Dependency in HRA." In: *2021 International Topical Meeting on Probabilistic Safety Assessment and Analysis (PSA 2021)* (Nov. 7–12, 2021). Columbus, OH, Nov. 2021.

[8] Alan Kolaczkowski, John Forester, Erasmia Lois, and Susan Cooper. *NUREG-1792: Good Practices for Implementing Human Reliability Analysis (HRA)*. Tech. rep. Washington, DC: U.S. Nuclear Regulatory Commission, 2005.

[9] April M. Whaley, Ronald L. Boring, Harold S. Blackman, Patrick H. McCabe, and Bruce P. Hallbert. "Lessons learned from dependency usage in HERA: Implications for THERP-related HRA methods." In: *IEEE Conference on Human Factors and Power Plants* (2007), pp. 322–327. DOI: 10.1109/HFPP.2007.4413227.

[10] S. Lewis and S. Cooper. *NUREG-1921: EPRI/NRC-RES Fire Human Reliability Analysis Guidelines*. Tech. rep. Washington, DC: U.S. Nuclear Regulatory Commission, 2012.

[11] Nathan Siu, Martin Stutzke, Suzanne Dennis, and Donald Harrison. *NUREG-2201: Probabilistic Risk Assessment and Regulatory Decisionmaking: Some Frequently Asked Questions.* Tech. rep. U.S. NRC Office of Nuclear Regulatory Research, 2016.

[12] Harold S. Blackman and Ronald L. Boring. "Assessing dependency in SPAR-H: Some practical considerations." In: *Advances in Intelligent Systems and Computing* 589 (2018), pp. 13–19. DOI: 10.1007/978-3-319-60645-3_2.

[13] M. A. B. Alvarenga and P. F. Frutuoso e Melo. "A review of the cognitive basis for human reliability analysis." In: *Progress in Nuclear Energy* 117 (2019). DOI: 10.1016/j.pnucene.2019.103050.

[14] Torrey Mortenson and Ronald Boring. "Is Dependency in Human Reliability Analysis a Real Phenomenon? Refining the Dependency Concept through Research." In: *Advances in Safety Management and Human Performance: Proceedings of the AHFE 2021 Virtual Conferences on Safety Management and Human Factors, and Human Error, Reliability, Resilience, and Performance* (July 25–29, 2021). Virtual, 2021.

[15] Ronald Boring, Jooyoung Park, and Torrey Mortenson. "Formative vs. Summative Dependence in Human Reliability Analysis." In: *Advances in Safety Management and Human Performance*. Ed. by Pedro M Arezes and Ronald L Boring. July. Cham: Springer International Publishing, 2021, pp. 62–69. ISBN: 978-3-030-80288-2. DOI: 10.1007/978-3-030-80288-2.

[16] A. D. Swain and H. E. Guttmann. *NUREG/CR-1278: Handbook of Human Reliability Analysis with Emphasis on Nuclear Power Plant Applications*. Tech. rep. Washington, DC: U.S. Nuclear Regulatory Commission, 1983.

[17] Fatmagul Ibisoglu and Katrina M Groth. "Methods for Dependency between Human Failure Events in Human Reliability Analysis: An Overview of the State-of-the-Art." In: *Proceedings of the 29th European Safety and Reliability Conference*. 2019. DOI: 10.3850/981-973-0000-00-0.

[18] Vincent P. Paglioni and Katrina M. Groth. "Unified Definitions for Dependency in Quantitative Human Reliability Analysis." In: *Proceedings of the 30th European Safety and Reliability Conference and the 15th Probabilistic Safety Assessment and Management Conference* (Nov. 1–5, 2020). Virtual, 2020.

[19] IEEE Power and Energy Society. *IEEE Std 1082-2017: IEEE Guide for Incorporating Human Reliability Analysis into Probabilistic Risk Assessments for Nuclear Power Generating Stations and Other Nuclear Facilities*. New York City, 2017.

[20] American Society of Mechanical Engineers. *ASME-RA-2002: Standard for probabilistic risk assessment for nuclear power plant applications*. Tech. rep. American Society of Mechanical Engineers, 2000.

[21] Barbara Jean Bell and Alan D Swain. *NUREG/CR-2254: A Procedure for Conducting a Human Reliability Analysis for Nuclear Power Plants*. Tech. rep. May. Washington, DC: U.S. Nuclear Regulatory Commission, 1983.

[22] J. Forester, A. Kolaczkowski, E. Lois, and D. Kelly. *NUREG-1842: Evaluation of Human Reliability Analysis Methods Against Good Practices*. Tech. rep. Washington, DC: U.S Nuclear Regulatory Commission, 2006, p. 279.

[23] P. K. Samanta, J. N. O'Brien, and H. W. Morrison. *NUREG/CR-3837: Multiple-Sequence Failure Model: Evaluation of and Procedures for Human Error Dependency*. Tech. rep. May. Upton: Brookhaven National Laboratory, 1985.

[24] M. T. Barriere, J. Wreathall, S. E. Cooper, D. C. Bley, W. J. Luckas, and A. Ramey-Smith. *NUREG-6265: Multidisciplinary Framework for Human Reliability Analysis with an Application to Errors of Commission and Dependencies*. Tech. rep. Washington, DC: U.S. Nuclear Regulatory Commission, 1995.

[25] J. M. O'Hara, J. C. Higgins, W. S. Brown, R. Fink, J. Persensky, P. Lewis, J. Kramer, and A. Szabo. *NUREG/CR-6947: Human Factors Considerations with Respect to Emerging Technology in Nuclear Power Plants*. Tech. rep. Upton, NY: Brookhaven National Laboratory, 2008.

[26] Bruce Hallbert, Alan Kolaczkowski, and Erasmia Lois. *NUREG/CR-6949: The employment of empirical data and Bayesian methods in human reliability analysis: a feasibility study*. Tech. rep. Idaho Falls, ID: Idaho National Laboratory, 2007, p. 85.

[27] Ronald Boring and David Gertman. *P-203: Human Reliability Analysis (HRA) Training Course*. Tech. rep. Washington, D.C.: U.S. Nuclear Regulatory Commission, 2016.

[28] Mohammad Modarres, Mark Kaminskiy, and Vasiliy Krivtsov. *Reliability Engineering and Risk Analysis: A Practical Guide*. 2nd. Taylor & Francis Group: CRC Press, 2009. ISBN: 978-0-8493-9247-4.

[29] Katrina M. Groth and Ali Mosleh. "Deriving causal Bayesian networks from human reliability analysis data: A methodology and example model." In: *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability* 226.4 (2012), pp. 361–379. DOI: 10.1177/1748006X11428107.

[30] Katrina M. Groth, Reuel Smith, and Ramin Moradi. "A hybrid algorithm for developing third generation HRA methods using simulator data, causal models, and cognitive science." In: *Reliability Engineering and System Safety* 191 (2019), p. 106507. DOI: 10.1016/j.ress.2019.106507.

[31] D. Gertman, H. Blackman, J. Marble, J. Byers, and C. Smith. *NUREG/CR-6883: The SPAR-H Human Reliability Analysis Method*. Tech. rep. Idaho Falls, ID: Idaho National Laboratory, 2005.

[32] Marko Čepin. "DEPEND-HRA-A method for consideration of dependency in human reliability analysis." In: *Reliability Engineering and System Safety* 93 (2008), pp. 1452–1460. DOI: 10.1016/j.ress.2007.10.004.

[33] Marko Čepin. "Comparison of methods for dependency determination between human failure events within human reliability analysis." In: *Science and Technology of Nuclear Installations* (2008). DOI: 10.1155/2008/987165.

[34] Marko Čepin. "Importance of human contribution within the human reliability analysis (IJS-HRA)." In: *Journal of Loss Prevention in the Process Industries* 21.3 (2008), pp. 268–276. DOI: 10.1016/j.jlp.2007.04.012.

[35] Massimiliano De Ambroggi and Paolo Trucco. "Modelling and assessment of dependent performance shaping factors through Analytic Network Process." In: *Reliability Engineering and System Safety* 96.7 (2011), pp. 849–860. DOI: 10.1016/j.ress.2011.03.004.

[36] Ronald Boring, Martin Rasmussen, Thomas Ulrich, Sarah Ewing, and Diego Mandelli. "Task and procedure level primitives for modeling human error." In: *International Conference on Applied Human Factors and Ergonomics, AHFE 2017* (July 17–21, 2017). Los Angeles, CA, 2017, pp. 30–40.

[37] R. L. Boring. "A dynamic approach to modeling dependence between Human Failure Events." In: *Safety and Reliability of Complex Engineered Systems - Proceedings of the 25th European Safety and Reliability Conference, ESREL 2015* (2015), pp. 2845–2851. DOI: 10.1201/b19094-373.

[38] Ronald L. Boring and Jeffrey C. Joe. "Task Decomposition in Human Reliability Analysis." In: *Probabilistic Safety Assessment and Management* (June 22–27, 2014). Honolulu, HI.

[39] Ronald L. Boring. "Fifty years of THERP and human reliability analysis." In: *11th International Probabilistic Safety Assessment and Management Conference and the Annual European Safety and Reliability Conference 2012, PSAM11 ESREL 2012* 5 (2012), pp. 3523–3532.

[40] Sarah M. Herberger and Ronald L. Boring. "Human Failure Event Dependence: What Are the Limits?" In: *13th International Conference on Probabilistic Safety Assessment and Management (PSAM 13)* (Oct. 2–7, 2016). Seoul.

[41] Jooyoung Park, Wondea Jung, and Jonghyun Kim. "Inter-relationships between performance shaping factors for human reliability analysis of nuclear power plants." In: *Nuclear Engineering and Technology* 52.1 (2020), pp. 87–100. DOI: 10.1016/j.net.2019.07.004.

[42] Katrina M Groth and Ali Mosleh. "A data-informed PIF hierarchy for model-based Human Reliability Analysis." In: *Reliability Engineering and System Safety* 108 (2012), pp. 154–174.

[43] Norman Fenton and Martin Neil. *Risk Assessment and Decision Analysis with Bayesian Networks*. 2nd. Boca raton: CRC Press, Taylor & Francis Group, 2018.

[44] Vincent P. Paglioni and Katrina M. Groth. "Dependency definitions for quantitative human reliability analysis." In: *Reliability Engineering & System Safety* (2022), p. 108274. DOI: https://doi.org/10.1016/j.ress.2021.108274.

[45] J. Xing, Y. J. Chang, and J. DeJesus Segarra. *NUREG-2198: The General Methodology of an Integrated Human Event Analysis System (IDHEAS-G)*. Tech. rep. Washington, D.C.: U.S. Nuclear Regulatory Commission, 2020.

[46] U.S. Nuclear Regulatory Commission. *WASH-1400: Reactor Safety Study - An Assessment of Accident Risks in U.S. Commercial Nuclear Power Plants*. Tech. rep. 1975.

[47] Huafei Liao, John Forester, Vinh N Dang, Andreas Bye, Yung Hsien J Chang, and Erasmia Lois. "Assessment of HRA method predictions against operating crew performance: Part II: Overall simulator data, HRA method predictions, and intra-method comparisons." In: *Reliability Engineering & System Safety* 191 (2019), p. 106510.

[48] S. E. Cooper, A. M. Ramey-Smith, J. Wreathall, G. W. Parry, D. C. Bley, W. J. Luckas, J. H. Taylor, and M. T. Barriere. *NUREG-6350: A Technique for Human Error Analysis (ATHEANA)*. Tech. rep. U.S. Nuclear Regulatory Commission, 1996.

[49] Michelle Kichline et al. *RIL 2021-14 - Integrated Human Event Analysis System Dependency Analysis Guidance*. U.S. Nuclear Regulatory Commission, 2021.

[50] Barry Kirwan. "The validation of three human reliability quantification techniques THERP, HEART and JHEDI: Part 1 - Technique descriptions and validation issues." In: *Applied Ergonomics* 27.6 (1996), pp. 359–373. DOI: 10.1016/S0003-6870(96)00044-0.

[51] Alan D. Swain. *NUREG-4772: Accident Sequence Evaluation Program Human Reliability Analysis Procedure*. Tech. rep. Albuquerque: U.S. Nuclear Regulatory Commission, 1987.

[52] Peng Cheng Li, Guo Hua Chen, Li Cao Dai, Li Zhang, Ming Zhao, and Wei Chen. "Methodology for analyzing the dependencies between human operators in digital control systems." In: *Fuzzy Sets and Systems* 293 (2016), pp. 127–143. DOI: 10.1016/j.fss.2015.04.002.

[53] Luca Podofillini, Vinh Dang, Enrico Zio, Piero Baraldi, and Massimo Librizzi. "Using expert models in human reliability analysis-a dependence assessment method based on fuzzy logic." In: *Risk Analysis* 30.8 (2010), pp. 1277–1297. DOI: 10.1111/j.1539-6924.2010.01425.x.

[54] Marzio Marseguerra, Enrico Zio, and Massimo Librizzi. "Human reliability analysis by fuzzy "CREAM"." In: *Risk Analysis* 27.1 (2007), pp. 137–154. DOI: 10.1111/j.1539-6924.2006.00865.x.

[55] E. Zio, P. Baraldi, M. Librizzi, L. Podofillini, and V. N. Dang. "A fuzzy set-based approach for modeling dependence among human errors." In: *Fuzzy Sets and Systems* 160.13 (2009), pp. 1947–1964. DOI: 10.1016/j.fss.2009.01.016.

[56] Xiaoyan Su, Sankaran Mahadevan, Peida Xu, and Yong Deng. "Dependence Assessment in Human Reliability Analysis Using Evidence Theory and AHP." In: *Risk Analysis* 35.7 (2015), pp. 1296–1316. DOI: 10.1111/risa.12347.

[57] Xianglin Zheng and Yong Deng. "Dependence assessment in human reliability analysis based on evidence credibility decay model and IOWA operator." In: *Annals of Nuclear Energy* 112 (2018), pp. 673–684. DOI: 10.1016/j.anucene.2017.10.045.

[58] Xingfeng Guo, Yanhui Zhou, Jin Qian, and Yong Deng. "Using Evidence Credibility Decay Model for dependence assessment in human reliability analysis." In: *Annals of Nuclear Energy* 100 (2017), pp. 107–118. DOI: 10.1016/j.anucene.2016.10.007.

[59] Luyuan Chen, Xinyi Zhou, Fuyuan Xiao, Yong Deng, and Sankaran Mahadevan. "Evidential Analytic Hierarchy Process Dependence Assessment Methodology in Human Reliability Analysis." In: *Nuclear Engineering and Technology* 49.1 (2017), pp. 123–133. DOI: 10.1016/j.net.2016.10.003.

[60] Xinyi Zhou, Xinyang Deng, Yong Deng, and Sankaran Mahadevan. "Dependence assessment in human reliability analysis based on D numbers and AHP." In: *Nuclear Engineering and Design* 313 (2017), pp. 243–252. DOI: 10.1016/j.nucengdes.2016.12.001.

[61] Zhi Yu Sun, Jian Lan Zhou, and Lin Fei Gan. "Safety assessment in oil drilling work system based on empirical study and Analytic Network Process." In: *Safety Science* 105.February (2018), pp. 86–97. DOI: 10.1016/j.ssci.2018.02.004.

[62] Jeffrey W. Herrmann. *Engineering Decision Making and Risk Management*. 1st ed. Hoboken: John Wiley & Sons, 2015. ISBN: 978-1-118-91933-0.

[63] Katrina M. Groth and Laura P. Swiler. "Bridging the gap between HRA research and HRA practice: A Bayesian network version of SPAR-H." In: *Reliability Engineering and System Safety* 115 (2013), pp. 33–42. DOI: 10.1016/j.ress.2013.02.015.

[64] Kilian Zwirglmaier, Daniel Straub, and Katrina M. Groth. "Capturing cognitive causal paths in human reliability analysis with Bayesian network models." In: *Reliability Engineering and System Safety* 158 (2017), pp. 117–129. DOI: 10.1016/j.ress.2016.10.010.

[65] Andres Ruiz-Tagle, Vincent P Paglioni, Enrique Lopez-Droguett, and Katrina M. Groth. "A Framework to Extrapolate and Evaluate Human Reliability Causal Models from Event Report Narratives." In: *2021 International Topical Meeting on Probabilistic Safety Assessment and Analysis (PSA 2021)*. Columbus, 2021.

[66] Tae Jin Kim and Poong Hyun Seong. "Influencing factors on situation assessment of human operators in unexpected plant conditions." In: *Annals of Nuclear Energy* (2019). DOI: 10.1016/j.anucene.2019.06.051.

[67] Jooyoung Park, Awwal Mohammed Arigi, and Jonghyun Kim. "A comparison of the quantification aspects of human reliability analysis methods in nuclear power plants." In: *Annals of Nuclear Energy* (2019). DOI: 10.1016/j.anucene.2019.05.031.

[68] Petek Yontay and Rong Pan. "A computational Bayesian approach to dependency assessment in system reliability." In: *Reliability Engineering and System Safety* 152 (2016), pp. 104–114. DOI: 10.1016/j.ress.2016.03.005.

[69] Huafei Liao, John Forester, Vinh N Dang, Andreas Bye, Yung Hsien J Chang, and Erasmia Lois. "Assessment of HRA method predictions against operating crew performance: Part I: Study background, design and methodology." In: *Reliability Engineering & System Safety* 191 (2019), p. 106509.

[70] Huafei Liao, John Forester, Vinh N Dang, Andreas Bye, Yung Hsien J Chang, and Erasmia Lois. "Assessment of HRA method predictions against operating crew performance: Part III: Conclusions and achievements." In: *Reliability Engineering & System Safety* 191 (2019), p. 106511.

[71] Jooyoung Park, Ronald L. Boring, and Jonghyun Kim. "An Identification of PSF Lag and Linger Effects for Dynamic Human Reliability Analysis: Application of Experimental Data." In: *International Conference on Human System Interaction, HSI* 2019-June (2019), pp. 12–16. DOI: 10.1109/HSI47298.2019.8942621.

[72] Jooyoung Park, Awwal Mohammed Arigi, and Jonghyun Kim. "Treatment of human and organizational factors for multi-unit HRA: Application of SPAR-H method." In: *Annals of Nuclear Energy* 132 (2019), pp. 656–678. DOI: 10.1016/j.anucene.2019.06.053.

[73] Finn V Jensen and Thomas Dyhre Nielsen. *Bayesian Networks and Decision Graphs*. New York City: Springer, 2007. ISBN: 0387682813.

[74] Judea Pearl. *Causality*. 2nd ed. Cambridge, UK: Cambridge University Press, 2009. ISBN: 978-0-521-89560-6. DOI: 10.1017/CBO9780511803161.

[75] Ross Shachter. "Bayes-Ball: The Rational Pastime (for Determining Irrelevance and Requisite Information in Belief Networks and Influence Diagrams)." In: *UAI'98: Proceedings of the Fourteenth Conference on Uncertainty in Artificial Intelligence* (July 24–26, 1998). Madison, WI, pp. 480–487.

[76] BayesFusion LLC. *GeNIe*. https://www.bayesfusion.com/genie/. 2015.

[77] L. Mkrtchyan, L. Podofillini, and V. N. Dang. "Bayesian belief networks for human reliability analysis: A review of applications and gaps." In: *Reliability Engineering and System Safety* 139 (2015), pp. 1–16. DOI: 10.1016/j.ress.2015.02.006.

[78] Piero Baraldi, Luca Podofillini, Lusine Mkrtchyan, Enrico Zio, and Vinh N. Dang. "Comparing the treatment of uncertainty in Bayesian networks and fuzzy expert systems used for a human reliability analysis application." In: *Reliability Engineering and System Safety* 138 (2015), pp. 176–193. DOI: 10.1016/j.ress.2015.01.016.

[79] L. Podofillini and V. N. Dang. "A Bayesian approach to treat expert-elicited probabilities in human reliability analysis model construction." In: *Reliability Engineering and System Safety* 117 (2013), pp. 52–64. DOI: 10.1016/j.ress.2013.03.015.

[80] Salvatore F Greco, Luca Podofillini, and Vinh N Dang. "A Bayesian model to treat within-category and crew-to-crew variability in simulator data for Human Reliability Analysis." In: *Reliability Engineering & System Safety* 206 (2021), p. 107309.

[81] J. C. Williams. "Validation of human reliability assessment techniques." In: *Reliability Engineering* 11 (1985), pp. 149–162. DOI: 10.1016/S0925-7535(97)00050-7.

[82] Nsimah J. Ekanem, Ali Mosleh, and Song Hua Shen. "Phoenix - A model-based Human Reliability Analysis methodology: Qualitative Analysis Procedure." In: *Reliability Engineering and System Safety* 145 (2016), pp. 301–315. DOI: 10.1016/j.ress.2015.07.009.

[83] Ronald Boring, Diego Mandelli, Jeffrey Joe, Curtis Smith, and Katrina Groth. *Light Water Reactor Sustainability Program: A Research Roadmap for Computation-Based Human Reliability Analysis*. Tech. rep. April. DOE-NE Light Water Reactor Sustainability Program, 2015.

[84] Y. James Chang et al. "The SACADA database for human reliability and human performance." In: *Reliability Engineering and System Safety* 125 (2014), pp. 117–133. DOI: 10.1016/j.ress.2013.07.014.

[85] Erik Hollnagel. *Cognitive reliability and error analysis method (CREAM)*. Elsevier, 1998.

[86] Shokoufeh Abrishami and Pieter Van Gelder. "Improving the performance of Success Likelihood Index Model (SLIM) using Bayesian Network." In: *Proceedings of the 29th European Safety and Reliability Conference*. 2019, pp. 978–981. ISBN: 9789811127243. DOI: 10.3850/978-981-11-2724-3.

[87] Ronald Boring, Diego Mandelli, Martin Rasmussen, Sarah Herberger, Thomas Ulrich, Katrina Groth, and Curtis Smith. "Human unimodel for nuclear technology to enhance reliability (HUNTER): A framework for computational-based human reliability analysis." In: *13th International Conference on Probabilistic Safety Assessment and Management (PSAM 13)*. Seoul, 2016.

[88] Vincent Philip Paglioni, Torrey Mortenson, and Katrina M. Groth. "The human failure event: what is it and what should it be?" In: *Proceedings of the 16th Probabilistic Safety Assessment and Management Conference (PSAM16)* (June 26–July 1, 2022). Honolulu, HI, 2022.

[89] Morten Lind. "MODELING GOALS AND FUNCTIONS OF COMPLEX INDUSTRIAL PLANTS." In: *Applied Artificial Intelligence* 8.2 (1994), pp. 259–283. DOI: 10.1080/08839519408945442.

[90] Mohammad Modarres. "A Functional Classification Based on Conservation Principles." In: *Proceedings of the Fifth International Functional Modeling Workshop* (July 1–3, 1997). Paris-Troyes, France, 1997.

[91] Atoosa Jalashgar. "Function-Oriented System Analysis: Putting the GTST, MFM and HMG Methods into Perspective." In: *Proceedings of the Fifth International Functional Modeling Workshop* (July 1–3, 1997). Paris-Troyes, France, 1997.

[92] Christine M. Mitchell. "The Importance of Modeling Operator Functions in Complex Dynamic Systems: A Conceptual Overview and Proposed Methodology." In: *Proceedings of the Third International Functional Modeling Workshop* (June 30–July 1, 1995). College Park, MD, 1995.

[93] Atoosa Jalashgar. "Modeling the Hierarchy of Basic Ingredients of Complex Human-Made Systems." In: *Proceedings of the Fourth International Functional Modeling Workshop* (May 19–20, 1996). Athens, Greece, 1996.

[94] Stuart K. Card, Thomas P. Moran, and Alan Newell. *The Psychology of Human-Computer Interaction*. CRC Press, 2017. ISBN: 9781138432956.

[95] Austin D. Lewis and Katrina M. Groth. "A Review of Methods for Discretizing Continuous-Time Accident Sequences." In: *Proceedings of the 29th European Safety and Reliability Conference (ESREL 19)* (Sept. 22–26, 2019). Hannover, Germany, 2019.

[96] Michelle Bensi, Armen Der Kiureghian, and Daniel Straub. "Efficient Bayesian network modeling of systems." In: *Reliability Engineering and System Safety* 112 (2013), pp. 200–213. DOI: 10.1016/j.ress.2012.11.017.

[97] David Lopez-Paz. "From Dependence to Causation." PhD thesis. University of Cambridge, 2016. arXiv: 1607.03300.

[98] Yochan Kim, Jinkyun Park, Wondea Jung, Sun Yeong Choi, and Seunghwan Kim. "Estimating the quantitative relation between PSFs and HEPs from full-scope simulator data." In: *Reliability Engineering & System Safety* 173 (2018), pp. 12–22.

[99] Ar Ryum Kim, Jinkyun Park, Yochan Kim, Jaewhan Kim, and Poong Hyun Seong. "Quantification of performance shaping factors (PSFs)' weightings for human reliability analysis (HRA) of low power and shutdown (LPSD) operations." In: *Annals of Nuclear Energy* 101 (2017), pp. 375–382. DOI: 10.1016/j.anucene.2016.11.009.

[100] Christopher M. Bishop. *Pattern Recognition and Machine Learning*. Berlin: Springer-Verlag, 2006. ISBN: 978-0-387-31073-2.

[101] Christopher Hitchcock. "Probabilistic Causation." In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Spring 2021. Metaphysics Research Lab, Stanford University, 2021.

[102] Wondea Jung, Jinkyun Park, Yochan Kim, Sun Yeong Choi, and Seunghwan Kim. "HuREX – A framework of HRA data collection from simulators in nuclear power plants." In: *Reliability Engineering and System Safety* 194 (2020), p. 106235. DOI: 10.1016/j.ress.2018.07.036.

[103] Yung Hsien James Chang, Yochan Kim, Jinkyun Park, and Lawrence Criscione. "SACADA and HuREX: Part 1. the use of SACADA and HuREX systems to collect human reliability data." In: *Nuclear Engineering and Technology* 54 (5 May 2022), pp. 1686–1697. DOI: 10.1016/J.NET.2021.10.037.

[104] Yochan Kim, Yung Hsien James Chang, Jinkyun Park, and Lawrence Criscione. "SACADA and HuREX part 2: The use of SACADA and HuREX data to estimate human error probabilities." In: *Nuclear Engineering and Technology* 54 (3 Mar. 2022), pp. 896–908. DOI: 10.1016/J.NET.2021.09.003.

[105] James Reason. *Human Error*. 1st. Cambridge: Cambridge University Press, 1990. ISBN: 978-0-521-31419-0.

[106] Yochan Kim, Jinkyun Park, and Mary Presley. "Hybrid Variable Selection for Estimating Nominal HEPs and PIF Multipliers for Digital Control Rooms." In: *Nuclear Science & Engineering* (2022). DOI: 10.1080/00295639.2022.2118481.

[107] Yochan Kim, Sun Yeong Choi, Jinkyun Park, and Jaewhan Kim. "Empirical study on human error probability of procedure-extraneous behaviors." In: *Reliability Engineering & System Safety* 227 (Nov. 2022), p. 108727. DOI: 10.1016/J.RESS.2022.108727.

[108] Yochan Kim and Jinkyun Park. "Incorporating prior knowledge with simulation data to estimate PSF multipliers using Bayesian logistic regression." In: *Reliability Engineering and System Safety* (2019). DOI: 10.1016/j.ress.2019.04.022.

[109] Yochan Kim. "Considerations for generating meaningful HRA data: Lessons learned from HuREX data collection." In: *Nuclear Engineering and Technology* (2020). DOI: 10.1016/j.net.2020.01.034.

[110] Y. H. J. Chang and A. Mosleh. "Cognitive modeling and dynamic probabilistic simulation of operating crew response to complex system accidents: Part 1: Overview of the IDAC Model." In: *Reliability Engineering & System Safety* 92 (8 Aug. 2007), pp. 997–1013. DOI: 10.1016/J.RESS.2006.05.014.

[111] Y. H. J. Chang and A. Mosleh. "Cognitive modeling and dynamic probabilistic simulation of operating crew response to complex system accidents. Part 2: IDAC performance influencing factors model." In: *Reliability Engineering and System Safety* 92.8 (2007), pp. 1014–1040. DOI: 10.1016/j.ress.2006.05.010.

[112] Y. H. J. Chang and A. Mosleh. "Cognitive modeling and dynamic probabilistic simulation of operating crew response to complex system accidents. Part 3: IDAC operator response model." In: *Reliability Engineering and System Safety* 92 (8 2007), pp. 1041–1060. DOI: 10.1016/j.ress.2006.05.013.

[113] Y. H. J. Chang and A. Mosleh. "Cognitive modeling and dynamic probabilistic simulation of operating crew response to complex system accidents. Part 4: IDAC causal model of operator problem-solving response." In: *Reliability Engineering and System Safety* 92 (8 2007), pp. 1061–1075. DOI: 10.1016/j.ress.2006.05.011.

[114] Y. H. J. Chang and A. Mosleh. "Cognitive modeling and dynamic probabilistic simulation of operating crew response to complex system accidents. Part 5: Dynamic probabilistic simulation of the IDAC model." In: *Reliability Engineering and System Safety* 92 (8 Aug. 2007), pp. 1076–1101. DOI: 10.1016/J.RESS.2006.05.012.

[115] A. P. Dempster, N. M. Laird, and D. B. Rubin. "Maximum Likelihood from Incomplete Data Via the EM Algorithm." In: *Journal of the Royal Statistical Society: Series B (Methodological)* 39.1 (1977), pp. 1–22. DOI: 10.1111/j.2517-6161.1977.tb01600.x.

[116] Charles R. Harris et al. "Array programming with NumPy." In: *Nature* 585.7825 (Sept. 2020), pp. 357–362. DOI: 10.1038/s41586-020-2649-2.

[117] Parot Ratnapinda and Marek J Druzdzel. "Learning discrete Bayesian network parameters from continuous data streams: What is the best strategy?" In: *Journal of Applied Logic* 13 (4, Part 2 2015). Special JAL Issue dedicated to Uncertain Reasoning at FLAIRS, pp. 628–642. DOI: https://doi.org/10.1016/j.jal.2015.03.007.

[118] Katrina M. Groth, Curtis L. Smith, and Laura P. Swiler. "A Bayesian method for using simulator data to enhance human error probabilities assigned by existing HRA methods." In: *Reliability Engineering and System Safety* 128 (2014), pp. 32–40. DOI: 10.1016/j.ress.2014.03.010.

[119] Marek Jozef Druzdzel and Linda C. van der Gaag. "Building probabilistic networks: "Where do the numbers come from?" guest editors' introduction." In: *IEEE Transactions on Knowledge and Data Engineering* 12 (4 July 2000), pp. 481–486. DOI: 10.1109/TKDE.2000.868901.

[120] Y. James Chang, Carmen Franklin, and Laurence Criscione. "Example Use of the SACADA Data to Inform HRA." In: *Proceedings of the 13th International Conference on Probabilistic Safety Assessment and Management (PSAM 13)* (Oct. 2–7, 2016). Seoul.

[121] Ahmad Al-Douri, Camille S. Levine, and Katrina M. Groth. "Identifying Human Failure Events (HFEs) for External Hazard Probabilistic Risk Assessment." In: *Reliability Engineering & System Safety* (2023), p. 109236. DOI: https://doi.org/10.1016/j.ress.2023.109236.

[122] Camille S. Levine, Ahmad Al-Douri, and Katrina Groth. "Causal Pathways Leading to Human Failure Events in Information-Gathering System Response Activities." In: *Proceedings of the 13th Nuclear Plant Instrumentation, Control, & Human-Machine Interface Technologies (NPIC&HMIT2023) Conference* (July 15–20, 2023). Knoxville, TN, 2023.

[123] John L Faundeen, Thomas E Burley, Jennifer Carlino, David L Govoni, Heather S Henkel, Sally Holl, Vivian B Hutchison, Elizabeth Martín, Ellyn T Montgomery, Cassandra C Ladino, et al. *The United States geological survey science data lifecycle model*. US Department of the Interior, US Geological Survey Reston, VA, USA, 2013.

[124] Michelle Bensi, Katrina Groth, Zeyun Wu, Zhegang Ma, Ray Schneider, Ahmad Al-Douri, Somayeh Mohammadi, Tao Liu, and Camille Levine. "Identifying and Prioritizing Sources of Uncertainty in External Hazard Probabilistic Risk Assessment: Project Activities and Progress." In: *Proceedings of the 16th Probabilistic Safety Assessment and Management Conference (PSAM 16)* (June 26–July 1, 2022). Honolulu, HI, 2022.