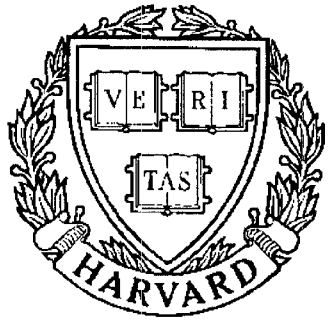
TECHNICAL RESEARCH REPORT

SYSTEMS RESEARCH CENTER



Supported by the National Science Foundation Engineering Research Center Program (NSFD CD 8803012), the University of Maryland, Harvard University, and Industry

Synthesis of Direct Mechanisms for Chemical Systems

by M.L. Mavrovouniotis

TR 91-44

SYNTHESIS OF DIRECT MECHANISMS FOR CHEMICAL SYSTEMS

Michael L. Mavrovouniotis
Department of Chemical Engineering and Systems Research Center
A.V. Williams Building
University of Maryland
College Park, MD 20742

ABSTRACT

A chemical system consists of intermediate species, terminal species, and mechanism steps. Understanding the behavior of a chemical system can be significantly aided by the identification of mechanisms responsible for overall reactions which do not involve net consumption or production of reaction intermediates. Issues arising in the definition and identification of direct mechanisms, which are the shortest possible mechanisms, are discussed. In the context of examples of catalytic synthesis of ammonia and methanol, an alternative approach for the construction of mechanisms from steps is presented. An algorithm for the construction of direct mechanism is then formally stated; the algorithm is based on successive processing and elimination of reaction intermediates which should not appear in the overall stoichiometry of the reactions accomplished by the mechanisms. Throughout the operation of the algorithm, irreversible steps are used only in their permitted direction. The basic algorithm may construct indirect or duplicate mechanisms, but variations of the algorithm are proposed which discard such redundant mechanisms. A number of hypothetical chemical systems illustrate the differences between the proposed algorithm and other approaches.

PART 1. PRESENTATION OF THE SYNTHESIS APPROACH IN THE CONTEXT OF SIMPLE EXAMPLES

INTRODUCTION

Consider a given set of elementary reaction steps which are feasible in a system, and the species involved in these steps. Following Happel and Sellers (1982, 1983, 1989), Happel (1986), Sellers (1984, 1989), and Happel *et al.* (1990), species can be classified as either *intermediates*, which occur in very small amounts, or *terminal* species which can occur in significant amounts and constitute the raw materials and products of reaction processes. An *overall* reaction, consisting of mechanism steps combined in specific proportions, involves the net transformation of only terminal species. Previous work on the theory of overall reaction mechanisms includes articles by Horiuti and Nakamura (1967), Horiuti (1973), and Temkin (1973, 1979).

The *synthesis of reaction mechanisms* involves the identification of sets of mechanism steps that accomplish such overall reactions involving terminal species. One is particularly interested in *direct* mechanisms (Milner, 1964), which are the smallest possible physically distinct mechanisms; direct mechanisms cannot be shortened through elimination of a step or reduced to a combination of smaller submechanisms. Once direct mechanisms are identified, all other possible mechanisms can be viewed as combinations of direct mechanisms. Another way to define direct mechanisms is to state that they are cycle-free, i.e., the steps participating in a direct mechanism cannot be combined into a "loop" or cycle accomplishing no net transformation; if there is a cycle in a mechanism then this mechanism contains excess steps and is not direct. What is usually referred to simply as a reaction mechanism in the literature is often implicitly required to be a direct mechanism, because physical intuition compels avoidance of cycles or excess steps. The concept of direct mechanisms has been well established by Happel and Sellers (1982, 1983, 1989), Happel (1986), Sellers (1984, 1989), and Happel *et al.* (1990), and will also be clearly set forth in this report, although under a somewhat different definition.

In considering the mechanism of a reaction, one often postulates certain steps and a particular mechanism which explains the reaction (and side-reactions leading to byproducts) in terms of the postulated steps. However, other mechanisms may also be possible within the postulated set of steps; the systematic generation and consideration of all possible direct mechanisms is valuable for a complete and sound analysis of a chemical system.

Sinanoğlu (1975), Sinanoğlu and Lee (1978), and Lee and Sinanoğlu (1981) introduced the concept of *laminar* mechanisms and presented procedures for the *a priori* (i.e., without regard to a specific set of mechanism steps) construction of reaction mechanisms given the number of catalysts involved. Milner (1964) showed that direct mechanisms for a given overall reaction can be constructed by examining combinations of I+1 steps at a time, where I is the number of intermediates, because a direct mechanism for a particular reaction can involve at most I+1 steps. Sellers (1971, 1972) also presented procedures for construction of mechanisms for certain chemical systems. Supported by a rigorous mathematical framework (Sellers, 1984, 1989), a more general and efficient algorithm has been presented by Happel and Sellers (1982, 1983, 1989) and Happel (1986) for the systematic construction of the direct mechanisms of any chemical system. A fundamental assumption for the operation of this algorithm is that all steps are reversible; the directionality of mechanism steps can be taken into account only at the end, by eliminating those overall reactions and mechanisms that use steps in a prohibited direction.

In this report, we present an alternative approach for the construction of reaction mechanisms, given a set of elementary steps and a classification of the participating species into intermediates and terminal species. The proposed algorithm is derived from a procedure for the synthesis of biochemical pathways from individual bioreactions (Mavrovouniotis, 1989, Mavrovouniotis *et al.*, 1990). The procedure for the synthesis of biochemical pathways clearly differs from the basic approach that will be presented in this report in its mathematical details; some conceptual differences will be addressed in the discussion section.

The report will follow, as closely as possible, the terminology and notation of Happel and Sellers (1982, 1983, 1989) and Happel (1986) except in those cases where the alternative algorithm introduced here requires a deviation in the notation, or particular complications in the notation of Happel and Sellers are

unnecessary for the analysis pursued in this report. We will assume that the reader is aware of the basic concepts and their technique. This report will summarize briefly only those ideas from Happel and Sellers (1982, 1983, 1989), Happel (1986), Sellers (1984, 1989), and Happel *et al.* (1990) that are most essential for the presentation and analysis of the new technique presented here.

It is hoped that the simplicity of the alternative algorithm and its potential efficiency will contribute to broader utilization of the concept of direct mechanisms. The scope of this report, however, does not include extensive argumentation in favor of using direct mechanisms and reactions in the analysis and modelling of reaction kinetics and mechanisms. A strong case for the physical significance of direct mechanisms and the value of generating them systematically has been made by Happel and Sellers (1982, 1983, 1989), Happel (1986), and Happel *et al.* (1990). The main goal of the first part of this report is the presentation of the basic features of the proposed alternative approach in the context of examples; the second part formalizes the proposed approach and demonstrates clearly that there are substantial differences between this algorithm and the previous work. The information in these first two articles is sufficient to allow application of the proposed method and establish its significance. Forthcoming publications will present additional refinements of the approach and large-scale applications which illustrate the importance of direct mechanisms.

BASIC NOTATION AND CONCEPTS

A chemical system consists of A species, designated as a_1, a_2, ..., a_A, and S mechanism steps, designated as s_1, s_2, ..., s_S. Each step s_i accomplishes a specific elementary reaction or transformation, $r_i=R(s_i)$. Let α_{ij} represents the stoichiometric coefficient of species a_j in step s_i, with the usual convention that $\alpha_{ij}>0$ if a_j is a product of s_i, $\alpha_{ij}<0$ if a_j is a reactant of s_i, and $\alpha_{ij}=0$ if a_j does not participate in s_i. The stoichiometry of the transformation $R(s_i)$ can then be written as:

$$r_i=R(s_i)=\sum_{j=1}^{A}\alpha_{ij}a_j$$

A chemical reaction, in the normal notation, can be obtained from $R(s_i)$ by setting it to zero and separating the terms into the two sides of the resulting equation:

$$\sum_{j\ (\alpha_{ij}<0)}(-\alpha_{ij})a_j \quad = \quad \sum_{j\ (\alpha_{ij}>0)}\alpha_{ij}a_j$$

The terms corresponding to a positive α_{ij} are products and appear in the right-hand side of the chemical equation, while the terms corresponding to a negative α_{ij} are reactants and appear in the left-hand side of the chemical equation; naturally, the terms for which $\alpha_{ij}=0$ vanish.

Because the directionality of reactions, mechanisms, and steps is important in the algorithm discussed in this report, the notation should be made clear in this regard. The above kind of chemical equation, with an equality sign (=) separating the two sides, will be used only when the issue of directionality is not taken into account at all. The sign $\leftrightarrows$ and the terms *reversible* or *bidirectional* will be used for a reaction, mechanism, or step whose net rate may be either positive (i.e., in the forward direction) or negative (i.e., in the reverse direction). Similarly, the sign $\rightarrow$ and the terms *irreversible* or *unidirectional* will denote a reaction, mechanism, or step which is either thermodynamically irreversible or known to proceed with a net positive rate (i.e., in the forward direction).

A *directionality label*, equal to either $\leftrightarrows$ or $\rightarrow$, can be assigned to each step in the chemical system. The directionality of a step s_i will be denoted by this directionality label preceding the step, i.e., as either $\leftrightarrows s_i$ or $\rightarrow s_i$.

The precise numerical values of the coefficients α_{ij} are physically significant if the steps are elementary; the coefficients then are integers and denote the order of the rate with respect to each reactant and product. However, since this report will not discuss specific rate expressions, the coefficients α_{ij} of each step s_i

can be considered meaningful only within an arbitrary multiplicative constant (a positive constant, if there are directionality connotations). Thus, the α_{ij} coefficients for each i will denote only ratios of species. The coefficients can, of course, be fixed by some convention, such as:

$$\sum_{j\ (\alpha_{ij}>0)} \alpha_{ij} = 1$$

A *mechanism*, m_k, is a combination of steps whose rates are restricted to be proportional to each other. In effect, if a given mechanism is the only one operating within a system, it is characterized by a single rate; this rate automatically determines the rates of all the steps, because their rates are proportional to the rate of the mechanism. Proportionality also applies to reaction extents. With σ_{ki} representing the proportions of the steps, a mechanism (or mechanistic vector) m_k is a linear combination of steps:

$$m_k=\sum_{i=1}^{S}\sigma_{ki}s_i$$

and belongs to an S-dimensional vector space. An additional restriction is imposed on the coefficients σ_{ki} to prevent violation of the proper direction of unidirectional steps. Unidirectional steps can only be used in the forward, i.e., with positive coefficients:

$$\rightarrow s_i \Rightarrow \sigma_{ki}\geq 0$$

As was the case with the coefficients α_{ij}, the coefficients σ_{ki} (for a given m_k) denote only proportions or ratios and are physically meaningful only within an arbitrary multiplicative constant, which, because of directionality considerations, must be positive. These coefficients can also be fixed by some additional convention, such as $\sum_{\sigma_{ki}>0} \sigma_{ki}= 1$ (for a particular mechanism m_k).

A reaction vector r_k, associated with each mechanism, describes the net transformation that is accomplished by the mechanism:

$$r_k=R(m_k)=\sum_{i=1}^{S}\sigma_{ki}R(s_i)$$

The use the symbols r_i and r_k for the reaction of a single step and a whole mechanism respectively, deviates from the notation of Happel and Sellers (1982, 1983, 1989) and Happel (1986). It will be seen that this is convenient for the alternative algorithm discussed here because the identification of reactions does not precede the identification of mechanisms, and because the individual steps will be initially treated as mechanisms. Substitution of the expression of $R(s_i)$ expands the expression for r_k into:

$$R(m_k)=\sum_{j=1}^{A} a_j \sum_{i=1}^{S}\sigma_{ki}\alpha_{ij}$$

If β_{kj} is the stoichiometric coefficient indicating the participation of species a_j in the net transformation accomplished by mechanism m_k,

$$\beta_{kj}=\sum_{i=1}^{S}\sigma_{ki}\alpha_{ij}$$

the reaction $R(m_k)$ can be written as:

$$R(m_k)=\sum_{j=1}^{A}\beta_{kj}a_j$$

Reaction r_k is said to involve a species a_j iff $\beta_{kj}\neq 0$. The remarks made previously on the transformations accomplished by individual steps, with respect to the conventions for the signs =, $\leftrightarrows$, and $\rightarrow$ apply to the net reactions of whole mechanisms as well. For clarity, the index i will generally be used for steps, the index j for species, and the index k for mechanisms.

A set of mechanisms is linearly independent iff their corresponding vectors of step coefficients (rows of σ_{ki}s) are linearly independent; similarly, reactions are linearly independent iff their species coefficients (rows of β_{kj}s) are linearly independent. The concept of linear independence is not relevant for the algorithm presented here, but plays a role in the algorithm of Happel and Sellers (1982, 1983, 1989) and Happel (1986).

The set of species a_j, j=1 to A, is divided into I intermediates a_j, j=1 to I, and T terminal species a_j, j=I+1 to I+T (where I+T=A); the Happel and Sellers (1982, 1983, 1989) convention of grouping intermediate species at the lower indices (j=1,..., I) and terminal species at the higher indices (j=I+1,..., I+T) is quite convenient and will be followed throughout this report.

Although intermediates are crucial for the operation of a reaction mechanism, they are present in very low concentrations. Over the course of a reaction, there is significant net production of some terminal species, and significant net consumption of others. For each intermediate, on the other hand, a high rate of production by some steps is balanced by a high rate of consumption by other steps, and the net rate is small during the normal progress of the reaction (excluding initiation and termination). This is just an informal way of stating the common pseudo-steady state assumption: During the normal progress of the reaction, the accumulation term in the mass balance for an intermediate can be set equal to zero, leaving only production and consumption rates in the mass balance. This explanation is sufficient, for our purposes, to attribute special significance to the interconversion of terminal species. A reaction that involves only terminal species is called an *overall reaction*, and a mechanism accomplishing an overall reaction is an *overall mechanism.* Mathematically, m_k is an overall mechanism and $r_k=R(m_k)$ an overall reaction, iff:

$$\beta_{kj}=0 \quad \text{for} \quad j=1 \text{ to } I$$

where β_{kj} was defined earlier as $\beta_{kj}=\sum_{i=1}^{S}\sigma_{ki}\alpha_{ij}$.

For a given set of steps, let U represent the set of overall mechanisms. A mechanism m_k is said to involve a step s_i iff $\sigma_{ki}\neq 0$. Let $T(m_k)$ denote the set of steps involved in a mechanism, and let the *length* of a mechanism be the cardinality of $T(m_k)$. An overall mechanism is a *direct mechanism* iff there is no other overall mechanism that involves only a subset of the steps involved in the original mechanism. In effect, an overall mechanism m_k is direct iff:

$$\nexists m_n\in U \ni T(m_n)\subset T(m_k)$$

Equivalently, every other overall mechanism must involve at least one step not present in the mechanism at hand, i.e., m_k is direct iff:

$$\forall m_n\in U: \exists s_i \ni [s_i\in T(m_k)\wedge s_i\notin T(m_n)]$$

Completely analogous to the definition of direct mechanisms in terms of σ_{ki} is the definition of direct reactions in terms of α_{ij} (Happel and Sellers, 1989), i.e., a reaction is direct if the set of species it involves is not a superset of the set of species of another overall reaction. Direct mechanisms do not necessarily correspond to direct reactions. There can exist mechanisms which are direct but accomplish overall reactions which are not direct; and, conversely, a direct reaction can have mechanisms which are not

direct. The algorithm that is presented in this report focuses on mechanisms and, because of this lack of correspondence between direct reactions and direct mechanisms, it does not need to consider direct reactions in any way.

The chemical significance of direct mechanisms has been firmly established by Happel and Sellers (1982, 1983, 1989), Happel (1986), and Happel *et al.* (1990). As they point out, what is usually referred to as a reaction mechanism in studies of chemical kinetics is implicitly a direct mechanism.

Direct mechanisms are not necessarily linearly independent; it may be possible to express a direct mechanism as a linear combination of other direct mechanisms. However, direct mechanisms are chemically distinct, because they involve different steps; for any pair of direct mechanisms, each one of them involves at least one step which is not present in the other. One may ask precisely when it is possible to obtain a mechanism which is direct and thus chemically significant, yet mathematically a linear combination of other direct mechanisms. The answer is evident in (and an important component of) the procedure of Happel and Sellers (1982, 1983, 1989), Happel (1986), and Sellers (1984); when two direct mechanisms involve the same step, a linear combination of them with coefficients that lead to elimination of the step may lead to another direct mechanism. The chemical significance of this combination mechanism is that it is feasible even if the step in question does not occur, and the rate of the mechanism is independent of the kinetics of that step; each of the mechanisms used in the combination, on the other hand, depends on the existence and rate of that step. Each direct mechanism corresponds to a unique combination steps; if, under some conditions, all other steps in the chemical system are disabled or very slow, then the direct mechanism in question will be the only feasible overall mechanism.

The definition of direct mechanisms given here differs in a subtle way from the definition given by Happel and Sellers (1982, 1983, 1989) and Sellers (1984). In particular, they define a mechanism m_k as *direct for a particular reaction r*, when no other mechanism for r involves a subset of the steps of m_k. In effect, a mechanism m_k is direct according to Happel and Sellers iff:

$$\nexists m_n \in U \ni \left(R(m_n)=R(m_k) \wedge T(m_n) \subset T(m_k) \right)$$

or:

$$\forall m_n \in U \ni R(m_n)=R(m_k) : \exists s_i \ni [s_i \in T(m_k) \wedge s_i \notin T(m_n)]$$

The definition used in this report, on the other hand, makes no particular reference to the reaction accomplished by the mechanism, other than the requirement that only overall mechanisms be considered. The difference is related to the operation of the algorithms. The Happel and Sellers (1982, 1983, 1989), Happel (1986), and Sellers (1984) algorithm constructs a basis set of reactions and then mechanisms for particular reactions, while the algorithm presented here proceeds directly to the construction of mechanisms, leaving the analysis of the reactions to be carried out, optionally, at the end. The consequences of this difference in the definitions is discussed in Part 2 in the context of examples.

Before proceeding with a description of the proposed alternative approach for the construction of direct mechanisms, let us note that the operation of the proposed algorithm differs significantly from the method of Happel and Sellers. First, the irreversibility of mechanism steps is taken into account by the proposed algorithm as mechanisms are constructed, i.e., it is not necessary to reserve directionality considerations until the end. Second, the identification of reactions does not precede the identification of mechanisms; reactions are not given much consideration in the operation of the algorithm. Third, direct mechanisms are constructed recursively from steps, rather than as combinations of mechanisms from a (basis) set of linearly independent mechanisms. A formal description and abstract examples in Part 2 further establish these differences. In the absence of formal analysis, or at least an empirical analysis based on a large and diverse set of examples, the evidence on the relative advantages of the two approaches and the cases in which one of the two might be easier or more efficient is inconclusive. There can be no doubt, however, that the two methods differ significantly.

AN EXAMPLE WITH REVERSIBLE STEPS

The basic features of the algorithm proposed here originated in an algorithm for the synthesis of biochemical pathways (Mavrovouniotis *et al.*, 1990), but, in the form in which it is presented here, the algorithm is specifically tailored to reaction mechanisms. One of the goals of this report is to demonstrate clearly that there are substantial differences between this algorithm and previous work.

The algorithm will first be presented in the context of an example on ammonia synthesis which combines the steps proposed by Horiuti (1973) and Temkin (1973). In this system, one assumes that all mechanism steps are reversible. The mechanism steps, given in Figure 1.1, are precisely those used by Happel and Sellers (1982, 1983) and Happel *et al.* (1990) (1990). The active site of the catalyst is denoted as "I" and all species that contain it are intermediates; N_2, H_2, and NH_3 are terminal species.

We start by considering mechanisms m_1, m_2, ..., m_9, each of which includes only the step with the same index, i.e.,

m_k: $\sigma_{ki}=\delta_{ki}$

where δ_{ki} is the Kronecker delta ($\delta_{ki}=1$ if $k=i$; $\delta_{ki}=0$ if $k\neq i$). Naturally, these are *not* overall mechanisms. A characteristic of the algorithm is that it operates on partial mechanisms, gradually transforming them to overall mechanisms. The partial mechanisms constructed above are all the mechanisms of length 1.

The initial setup for the application of the algorithm is shown in Figure 1.2. This setup, analogous to some used by Happel and Sellers (1982, 1983, 1989), lists for each mechanism m_k the quantities σ_{ki} and β_{kj}; entries that are equal to zero have been omitted from the two matrices. Note that intermediate species receive the lowest indices and are separated from terminal species. Instead of representing a mechanism as an explicit (algebraic) linear combination of steps, as Happel and Sellers do, Figure 1.2 shows the matrix of σ_{ki}s. This is merely a cosmetic difference; when there are few but densely connected steps, the setup shown here is probably more convenient; but when there is a large number of sparsely connected steps, it is more practical to replace the σ_{ki}s with an explicit linear combination of steps. Symbols indicating the reversibility of steps and mechanisms are also included in Figure 1.2. Directionalities are taken into account as the mechanisms are constructed and not only at the end; for this example, however, all steps are considered reversible and the directionality does not matter. The row labelled "number of combinations" and the column labelled "origin" will be explained in the description of the operation of the algorithm.

The algorithm proceeds by successively eliminating each intermediate species from the system. In order to eliminate an intermediate, we consider all the mechanisms whose reactions involve the intermediate species at hand. We can create combinations of two mechanisms at a time, with combination coefficients such that the intermediate vanishes from the reaction. If, for example, we pursue the elimination of NHI, there are four existing partial mechanisms, m_3, m_5, m_6, and m_7, to be considered; there are $4\times3/2=6$ pairwise combinations of these. For each pairwise combination (m_k, m_b), we can form a new mechanism as a linear combination of the two mechanisms; to ensure that the resultant mechanism no longer includes NHI (i.e., species a_3) in its net reaction, the new mechanism is formed using the coefficients of a_3 in the old mechanisms:

$\beta_{b3}\, m_k - \beta_{k3}\, m_b$

For example, if we combine m_3 and m_6 we can create the mechanism $2m_6+m_3$ (if we take $b=3$ and $k=6$), or $-m_3-2m_6$ (if $b=6$ and $k=3$); it does not matter which of the two possible assignments is chosen here, because the bidirectionality of the steps allows the two resulting mechanisms to be considered identical.

The row for a linear combination of mechanisms has elements σ_{ki} (left portion of Figure 1.2) and β_{kj} (right portion of Figure 1.2) which are simply the linear combinations of the respective elements of the old mechanisms.

Once we create all 6 combinations, we can eliminate the original mechanisms m_3, m_5, m_6, and m_7. To justify this transformation of the set of partial mechanisms, consider the set of mechanisms as a basis from which all other mechanisms (partial or overall) can be formed as linear combinations. All mechanisms which do not involve a_3 in their overall reaction can be constructed using the new basis which no longer includes m_3, m_5, m_6, and m_7, but includes their pairwise combinations that eliminate a_3. In effect, the old set was a basis for all mechanisms, while the new set of mechanisms is a basis for precisely those mechanisms that do not involve a_3 in their overall reaction.

The choice to eliminate the intermediate a_3 (NHI) first, however, is a poor one; it leads to construction of 6 new mechanisms (and abolition of 4 mechanisms). If we choose NH_2I (a_6) instead, we only need to

s_1: $N_2 + I \leftrightarrows N_2I$
s_2: $N_2I + H_2 \leftrightarrows N_2H_2I$
s_3: $N_2H_2I + I \leftrightarrows 2NHI$
s_4: $N_2 + 2I \leftrightarrows 2NI$
s_5: $NI + HI \leftrightarrows NHI + I$
s_6: $NHI + HI \leftrightarrows NH_2I$
s_7: $NHI + H_2 \leftrightarrows NH_3 + I$
s_8: $H_2 + I \leftrightarrows 2HI$
s_9: $NH_2I + HI \leftrightarrows NH_3 + I$

Figure 1.1. Mechanism steps for the synthesis of ammonia, as used in an example by Happel and Sellers (1982, 1983) and Happel *et al.* (1990); the steps were derived from mechanisms proposed by Horiuti (1973) and Temkin (1973). The symbol "I" denotes an active surface site on the catalyst.

$N_2+I \leftrightarrows N_2I$	$N_2I+H_2 \leftrightarrows N_2H_2I$	$N_2H_2I+I \leftrightarrows 2NHI$	$N_2+2I \leftrightarrows 2NI$	$NI+HI \leftrightarrows NHI+I$	$NHI+HI \leftrightarrows NH_2I$	$NHI+H_2 \leftrightarrows NH_3+I$	$H_2+I \leftrightarrows 2HI$	$NH_2I+HI \leftrightarrows NH_3+I$	mechanism ⇓	species ⇒	N_2I	N_2H_2I	NHI	NI	HI	NH_2I	I	N_2	H_2	NH_3
s_1	s_2	s_3	s_4	s_5	s_6	s_7	s_8	s_9		number of combinations ⇒	1	1	6	1	6	1	28	–	–	–
⇆	⇆	⇆	⇆	⇆	⇆	⇆	⇆	⇆		origin ⇓	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}
1									⇆ m_1	—	1						–1	–1		
	1								⇆ m_2	—	–1	1							–1	
		1							⇆ m_3	—		–1	2				–1			
			1						⇆ m_4	—				2			–2	–1		
				1					⇆ m_5	—			1	–1	–1		1			
					1				⇆ m_6	—			–1		–1	1	1			
						1			⇆ m_7	—			–1				1		–1	1
							1		⇆ m_8	—					2		–2		–1	
								1	⇆ m_9	—					–1	–1	2			1

Figure 1: Initial setup for the application of the algorithm on the ammonia mechanism. Entries that are equal to zero have been omitted from the matrices. The number of combination mechanisms that must be created to eliminate each species is listed below the species. Here, all steps are reversible; the number of combinations is simply x(x–1)/2, where x is the number of mechanisms whose reactions involve the species in question. Combinations are not shown for terminal species, because they will not be eliminated. The column for NH_2I (a_6) has been highlighted, because this species will be eliminated next.

$N_2 + I \leftrightarrows N_2I$	$N_2I + H_2 \leftrightarrows N_2H_2I$	$N_2H_2I + I \leftrightarrows 2NHI$	$N_2 + 2I \leftrightarrows 2NI$	$NI + HI \leftrightarrows NHI + I$	$NHI + HI \leftrightarrows NH_2I$	$NHI + H_2 \leftrightarrows NH_3 + I$	$H_2 + I \leftrightarrows 2HI$	$NH_2I + HI \leftrightarrows NH_3 + I$	mechanism	species ⇒	N_2I	N_2H_2I	NHI	NI	HI	NH_2I	—	N_2	H_2	NH_3
s_1	s_2	s_3	s_4	s_5	s_6	s_7	s_8	s_9	⇓	number of combinations ⇒	1	1	6	1	3	##	21	–	–	–
⇆	⇆	⇆	⇆	⇆	⇆	⇆	⇆	⇆		origin ⇓	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}
1									⇆ m_1	—	1					##	–1	–1		
	1								⇆ m_2	—	–1	1				##			–1	
		1							⇆ m_3	—		–1	2			##	–1			
			1						⇆ m_4	—				2		##	–2	–1		
				1					⇆ m_5	—			1	–1	–1	##	1			
##	##	##	##	##	##	##	##	##	⇆ m_6	—	##	##	##	##	##	##	##	##	##	##
						1			⇆ m_7	—			–1			##	1		–1	1
							1		⇆ m_8	—					2	##	–2		–1	
##	##	##	##	##	##	##	##	##	⇆ m_9	—	##	##	##	##	##	##	##	##	##	##
					1			1	⇆ m_{10}	a_6: $m_6 + m_9$			–1		–2	##	3			1

Figure 3. Setup for the application of the algorithm on the ammonia mechanism, after the elimination of a_6 (NH_2I). A new mechanism, m_{10}, has been constructed by combining m_6 and m_9 to eliminate a_6 from the net reaction. The numbers of combinations have been recomputed for those species participating in the abolished mechanisms; only the numbers for a_5 and a_7 have changed. The columns for a_2 (N_2H_2I) and a_4 (NI) have been highlighted; both of these species (which involve different mechanisms) will be eliminated next.

$N_2+I \leftrightarrows N_2I$	$N_2I+H_2 \leftrightarrows N_2H_2I$	$N_2H_2I+I \leftrightarrows 2NHI$	$N_2+2I \leftrightarrows 2NI$	$NI+HI \leftrightarrows NHI+I$	$NHI+HI \leftrightarrows NH_2I$	$NHI+H_2 \leftrightarrows NH_3+I$	$H_2+I \leftrightarrows 2HI$	$NH_2I+HI \leftrightarrows NH_3+I$	mechanism	species ⇒	N_2I	N_2H_2I	NHI	NI	HI	NH_2I	—	N_2	H_2	NH_3
s_1	s_2	s_3	s_4	s_5	s_6	s_7	s_8	s_9	⇓	number of combinations ⇒	1	##	6	##	3	##	10	–	–	–
⇆	⇆	⇆	⇆	⇆	⇆	⇆	⇆	⇆		origin ⇓	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}
1									⇆ m_1	—	1	##		##		##	–1	–1		
##	##	##	##	##	##	##	##	##	⇆ m_2	—	##	##	##	##	##	##	##	##	##	##
##	##	##	##	##	##	##	##	##	⇆ m_3	—	##	##	##	##	##	##	##	##	##	##
##	##	##	##	##	##	##	##	##	⇆ m_4	—	##	##	##	##	##	##	##	##	##	##
##	##	##	##	##	##	##	##	##	⇆ m_5	—	##	##	##	##	##	##	##	##	##	##
##	##	##	##	##	##	##	##	##	⇆ m_6	—	##	##	##	##	##	##	##	##	##	##
						1			⇆ m_7	—		##	–1	##		##	1		–1	1
							1		⇆ m_8	—		##		##	2	##	–2		–1	
##	##	##	##	##	##	##	##	##	⇆ m_9	—	##	##	##	##	##	##	##	##	##	##
					1			1	⇆ m_{10}	a_6: $m_6 + m_9$		##	–1	##	–2	##	3			1
	1	1							⇆ m_{11}	a_2: $m_2 + m_3$	–1	##	2	##		##	–1		–1	
			1	2					⇆ m_{12}	a_4: $m_4 + 2m_5$		##	2	##	–2	##		–1		

Figure 4. Application of the algorithm on the ammonia mechanism, after the elimination of a_2 and a_4. Two new mechanisms, m_{11} and m_{12}, have been constructed. The number of combinations has been recomputed for each remaining intermediate species. The intermediate N_2I (a_1) will be eliminated next.

$N_2+I \leftrightarrows N_2I$	$N_2I+H_2 \leftrightarrows N_2H_2I$	$N_2H_2I+I \leftrightarrows 2NHI$	$N_2+2I \leftrightarrows 2NI$	$NI+HI \leftrightarrows NHI+I$	$NHI+HI \leftrightarrows NH_2I$	$NHI+H_2 \leftrightarrows NH_3+I$	$H_2+I \leftrightarrows 2HI$	$NH_2I+HI \leftrightarrows NH_3+I$	mechanism	species ⇒	N_2I	N_2H_2I	NHI	NI	HI	NH_2I	I	N_2	H_2	NH_3
s_1	s_2	s_3	s_4	s_5	s_6	s_7	s_8	s_9	⇓	number of combinations ⇒	##	##	6	##	3	##	6	–	–	–
⇆	⇆	⇆	⇆	⇆	⇆	⇆	⇆	⇆		origin ⇓	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}
##	##	##	##	##	##	##	##	##	⇆ m_1	—	##	##	##	##	##	##	##	##	##	##
##	##	##	##	##	##	##	##	##	⇆ m_2	—	##	##	##	##	##	##	##	##	##	##
##	##	##	##	##	##	##	##	##	⇆ m_3	—	##	##	##	##	##	##	##	##	##	##
##	##	##	##	##	##	##	##	##	⇆ m_4	—	##	##	##	##	##	##	##	##	##	##
##	##	##	##	##	##	##	##	##	⇆ m_5	—	##	##	##	##	##	##	##	##	##	##
##	##	##	##	##	##	##	##	##	⇆ m_6	—	##	##	##	##	##	##	##	##	##	##
						1			⇆ m_7	—	##	##	–1	##		##	1		–1	1
							1		⇆ m_8	—	##	##		##	2	##	–2		–1	
##	##	##	##	##	##	##	##	##	⇆ m_9	—	##	##	##	##	##	##	##	##	##	##
					1			1	⇆ m_{10}	a_6: $m_6 + m_9$	##	##	–1	##	–2	##	3			1
##	##	##	##	##	##	##	##	##	⇆ m_{11}	a_2: $m_2 + m_3$	##	##	##	##	##	##	##	##	##	##
			1	2					⇆ m_{12}	a_4: $m_4 + 2m_5$	##	##	2	##	–2	##		–1		
1	1	1							⇆ m_{13}	a_1: $m_1 + m_{11}$	##	##	2	##		##	–2	–1	–1	

Figure 5. Application of the algorithm on the ammonia mechanism, after the elimination of a_1. Only 5 mechanisms, m_7, m_8, m_{10}, m_{12}, and m_{13}, survive at this point. The intermediate HI (a_5) will be eliminated next.

$N_2 + I \leftrightarrows N_2I$	$N_2I + H_2 \leftrightarrows N_2H_2I$	$N_2H_2I + I \leftrightarrows 2NHI$	$N_2 + 2I \leftrightarrows 2NI$	$NI + HI \leftrightarrows NHI + I$	$NHI + HI \leftrightarrows NH_2I$	$NHI + H_2 \leftrightarrows NH_3 + I$	$H_2 + I \leftrightarrows 2HI$	$NH_2I + HI \leftrightarrows NH_3 + I$	mechanism	species ⇒	N_2I	N_2H_2I	NHI	NI	HI	NH_2I	I	N_2	H_2	NH_3
s_1	s_2	s_3	s_4	s_5	s_6	s_7	s_8	s_9	⇓	number of combinations ⇒	##	##	10	##	##	##	10	–	–	–
⇆	⇆	⇆	⇆	⇆	⇆	⇆	⇆	⇆		origin ⇓	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}
						1			⇆ m_7	—	##	##	–1	##	##	##	1		–1	1
##	##	##	##	##	##	##	##	##	⇆ m_8	—	##	##	##	##	##	##	##	##	##	##
##	##	##	##	##	##	##	##	##	⇆ m_{10}	a_6: $m_6 + m_9$	##	##	##	##	##	##	##	##	##	##
##	##	##	##	##	##	##	##	##	⇆ m_{12}	a_4: $m_4 + 2m_5$	##	##	##	##	##	##	##	##	##	##
1	1	1							⇆ m_{13}	a_1: $m_1 + m_{11}$	##	##	2	##	##	##	–2	–1	–1	
					1		1	1	⇆ m_{14}	a_5: $m_8 + m_{10}$	##	##	–1	##	##	##	1		–1	1
			1	2			1		⇆ m_{15}	a_5: $m_8 + m_{12}$	##	##	2	##	##	##	–2	–1	–1	
			1	2	–1			–1	⇆ m_{16}	a_5: $m_{12} - m_{10}$	##	##	3	##	##	##	–3	–1		–1

Figure 6. Application of the algorithm on the ammonia mechanism, after the elimination of a_5. The mechanisms that had been crossed out in Figure 5 (i.e., mechanisms m_1, m_2, m_3, m_4, m_5, m_6, m_9, and m_{11}) have been removed altogether, to keep the figure small. Either of the remaining two intermediates (a_3 and a_7) can be eliminated next; a_3 is chosen.

$N_2 + I \leftrightarrows N_2I$	$N_2I + H_2 \leftrightarrows N_2H_2I$	$N_2H_2I + I \leftrightarrows 2NHI$	$N_2 + 2I \leftrightarrows 2NI$	$NI + HI \leftrightarrows NHI + I$	$NHI + HI \leftrightarrows NH_2I$	$NHI + H_2 \leftrightarrows NH_3 + I$	$H_2 + I \leftrightarrows 2HI$	$NH_2I + HI \leftrightarrows NH_3 + I$	mechanism ⇓	corresponding mechanism from Happel *et al.* (1990) ⇓	species ⇒	N_2I	N_2H_2I	NHI	NI	HI	NH_2I	I	N_2	H_2	NH_3
s_1	s_2	s_3	s_4	s_5	s_6	s_7	s_8	s_9			number of combinations ⇒	##	##	##	##	##	##	0	–	–	–
⇆	⇆	⇆	⇆	⇆	⇆	⇆	⇆	⇆			origin ⇓	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}
1	1	1				2			⇆ m_{17}	m_4,Table V	a_3: $m_{13} + 2m_7$	##	##	##	##	##	##		–1	–3	2
					1	–1	1	1	⇆ m_{18}	null,Table IV	a_3: $m_{14} - m_7$	##	##	##	##	##	##		0	0	0
			1	2		2	1		⇆ m_{19}	m_1,Table V	a_3: $m_{15} + 2m_7$	##	##	##	##	##	##		–1	–3	2
			1	2	–1			–1	⇆ m_{20}	m_3,Table V	a_3: $m_{16} + 3m_7$	##	##	##	##	##	##		–1	–3	2
1	1	1			2		2	2	⇆ m_{21}	m_5,Table V	a_3: $m_{13} + 2m_{14}$	##	##	##	##	##	##		–1	–3	2
–1	–1	–1	1	2			1		⇆ m_{22}	null,Table IV	a_3: $m_{15} - m_{13}$	##	##	##	##	##	##		0	0	0
3	3	3	–2	–4	2			2	⇆ m_{23}	m_6,Table V	a_3: $3m_{13} - 2m_{16}$	##	##	##	##	##	##		–1	–3	2
			1	2	2		3	2	⇆ m_{24}	m_2,Table V	a_3: $m_{15} + 2\,m_{14}$	##	##	##	##	##	##		–1	–3	2
			1	2	2		3	2	⇆ m_{25}	m_2,Table V	a_3: $m_{16} + 3m_{14}$	##	##	##	##	##	##		–1	–3	2
			1	2	2		3	2	⇆ m_{26}	m_2,Table V	a_3: $3m_{15} - 2m_{16}$	##	##	##	##	##	##		–1	–3	2

Figure 7. Application of the algorithm on the ammonia mechanism, after the elimination of a_3. Only the surviving mechanisms are shown. The intermediate a_7 does not need to be eliminated because its column contains only zeroes; this happened because, in Figure 6, the columns for a_3 and a_7 contained precisely opposite numbers and elimination of a_3 caused automatic elimination of a_7. An additional column shows the mechanisms from Happel *et al.* (1990) that are identical to the ones produced here. Note that the last three mechanisms are actually the same; there are only 8 distinct mechanisms, 2 of which (m_{18} and m_{22}) are mechanisms for the null overall reaction.

construct one mechanism (and abolish 2). Clearly, either choice is possible, but elimination of NH_2I keeps the number of active mechanisms lower. To permit the best choice of intermediate to be made, the setup of the algorithm (Figure 1.2) lists, below each species, the number of combination mechanisms that must be created to eliminate the species. Here, all steps are reversible; the number of combinations is simply $x(x-1)/2$, where x is the number of mechanisms whose reactions involve the species in question. Combinations are not shown for terminal species, because they will not be eliminated. The column for NH_2I (a_6) has been highlighted in Figure 1.2, to show that this is the species which will be eliminated first. Intuitively, one can think of this elimination as follows: There are only two steps that involve NH_2I; if any mechanism uses one of the two steps, it must use the other, with a coefficient that would eliminate the intermediate NH_2I; hence, the two steps can be combined.

When the elimination procedure is thus carried out on NH_2I (a_6), the setup shown in Figure 1.3 is obtained. In the "origin" column, the construction of the mechanism m_{10} is documented as resulting from the elimination of a_3 through a specific linear combination of the mechanisms m_6 and m_9. The row of m_{10} has been obtained by carrying out the linear combination of the rows of m_6 and m_9. The rows for m_6 and m_9, as well as the column for a_3 have been crossed out from the table. The numbers of combinations have changed for a_5 and a_7 (but remain unchanged for other intermediates). Any one of the species a_1, a_2, and a_4, each entailing only one new combination, could be eliminated next. Figure 1.3 shows both a_2 (N_2H_2I) and a_4 (NI) highlighted; their elimination can be carried out in parallel, because the two sets of mechanisms involved are disjoint, i.e., there is no mechanism whose reaction stoichiometry involves both a_4 and a_2.

Figure 1.4 reflects the state of the setup after the elimination of a_2 and a_4, which yielded two new partial mechanisms, m_{11} and m_{12}, and eliminated m_2, m_3, m_4, and m_5. The intermediate N_2I (a_1) is eliminated next, leading to the setup of Figure 1.5, in which only 5 mechanisms (m_7, m_8, m_{10}, m_{12}, and m_{13}) remain active. This reduction in the number of mechanisms as the first few intermediates are eliminated is quite common; however, the number of mechanisms tends to increase again at the end. Elimination of intermediate HI (a_5) next produces Figure 1.6. We can generally, at any time, drop inactive (crossed out) mechanisms or eliminated intermediates, but we have so far maintained them so that one can easily keep track of the progress of the algorithm. At this point, in order to reduce the size of the setup, we drop from the figure all those mechanisms that had been crossed out in Figure 1.5 or earlier.

The two intermediates (a_3 and a_7) remaining in Figure 1.6 involve the same number of combinations; elimination of either one results automatically in the elimination of the other (Figure 1.7), because in Figure 1.6 the columns for a_3 and a_7 contained precisely opposite numbers.

Figure 1.7 shows the final set of mechanisms produced here and their correspondence to the mechanisms constructed by Happel *et al.* (1990). All the direct mechanisms identified by Happel *et al.* (1990) have been produced. However, the last three mechanisms in Figure 1.7 are actually identical. Hence, the simple procedure discussed here does not preclude multiple occurrences of the same mechanism, or, as the next example will show, the occurrence of mechanisms that are not direct. For small studies this is not a significant drawback, because one can easily eliminate the redundancies in the end. The potential duplication of mechanisms and construction of indirect mechanisms are discussed and addressed in more detail in Part 2, showing how redundant mechanisms can be readily recognized and discarded.

AN EXAMPLE BASED ON IRREVERSIBLE STEPS

Consider, as a second example, the mechanism proposed by Yarlagadda *et al.* (1988) for the synthesis of methanol (Figure 1.8). This mechanism, analyzed by Happel *et al.* (1990), consists entirely of steps with a defined direction. Stepwise application of the algorithm proposed in this report, will show that the directionality of steps can be taken into account as the mechanisms are constructed.

In the beginning of their analysis, Happel *et al.* (1990) remove steps s_9 and s_{11}, because they lead to byproducts (ethane and dimethylether) rather than methanol. Here, we will keep all steps; it turns out that the products are not formed independently of each other, and isolation of the main product from the byproducts of the mechanism can be misleading. Happel *et al.* (1990) also initially consider all steps reversible and take into account the direction of the steps at the end. Since the algorithm that is proposed in this report can take directionality into account right from the start, the steps will be considered unidirectional (either irreversible or at least required to have a net rate in the forward direction). The species CH_4, O_2, CH_3OH, CO, H_2O, C_2H_6, and CH_3OCH_3 are terminal and all others are intermediates; formaldehyde (CH_2O) is an intermediate because it is present in small amounts (Happel *et al.*, 1990). The

s_1: $CH_4 + O_2 \rightarrow CH_3 + HO_2$

s_2: $CH_3 + O_2 \rightarrow CH_3O_2$

s_3: $CH_3O_2 \rightarrow CH_2O + OH$

s_4: $CH_3O_2 + CH_4 \rightarrow CH_3O_2H + CH_3$

s_5: $CH_3O_2H \rightarrow CH_3O + OH$

s_6: $CH_3O \rightarrow CH_2O + H$

s_7: $CH_3O + CH_4 \rightarrow CH_3OH + CH_3$

s_8: $OH + CH_4 \rightarrow CH_3 + H_2O$

s_9: $CH_3 + CH_3 \rightarrow C_2H_6$

s_{10}: $CH_3 + OH \rightarrow CH_3OH$

s_{11}: $CH_3 + CH_3O \rightarrow CH_3OCH_3$

s_{12}: $CH_2O + CH_3 \rightarrow CH_4 + CHO$

s_{13}: $CHO + O_2 \rightarrow CO + HO_2$

s_{14}: $CH_2O + CH_3O \rightarrow CH_3OH + CHO$

s_{15}: $CHO + CH_3 \rightarrow CO + CH_4$

Figure 1.8. Mechanism steps for the synthesis of methanol, as used in an example by Happel *et al.* (1990). The mechanism was proposed by Yarlagadda *et al.* (1988) and assumes that all steps have a net rate in the indicated direction.

s_1 →	s_2 →	s_3 →	s_4 →	s_5 →	s_6 →	s_7 →	s_8 →	s_9 →	s_{10} →	s_{11} →	s_{12} →	s_{13} →	s_{14} →	s_{15} →	mechanism ⇓	species ⇒	CH_3	HO_2	CH_3O_2	OH	CH_3O_2H	CH_3O	H	CHO	CH_2O
																combinations⇒	24	0	2	4	1	4	0	4	4
																origin ⇓	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9
1															→m_1	—	1	1							
	1														→m_2	—	−1		1						
		1													→m_3	—			−1	1					1
			1												→m_4	—	1		−1		1				
				1											→m_5	—				1	−1	1			
					1										→m_6	—						−1	1		1
						1									→m_7	—	1					−1			
							1								→m_8	—	1			−1					
								1							→m_9	—	−2								
									1						→m_{10}	—	−1			−1					
										1					→m_{11}	—	−1					−1			
											1				→m_{12}	—	−1							1	−1
												1			→m_{13}	—		1						−1	
													1		→m_{14}	—						−1		1	−1
														1	→m_{15}	—	−1							−1	

Figure 9 Initial setup for the application of the algorithm on the methanol mechanism. The number of combination mechanisms that must be created to eliminate each species is listed below the species; here, all steps are irreversible and the number of combinations is equal to yz, where z the number of mechanisms for which the species is a net product, and y is the number of mechanisms for which the species in question is a net reactant. Terminal species are not shown in this figure; they will be incorporated in Figure 10. The species a_2 (HO_2) and a_7 (H) will be eliminated first; they do not cause creation of any new mechanisms and they allow us to abolish m_1, m_6, and m_{13}.

s_1 →	s_2 →	s_3 →	s_4 →	s_5 →	s_6 →	s_7 →	s_8 →	s_9 →	s_{10} →	s_{11} →	s_{12} →	s_{13} →	s_{14} →	s_{15} →	mechanism ⇓	species ⇒	CH_3	CH_3O_2	OH	CH_3O_2H	CH_3O	CHO	CH_2O	CH_4	O_2	CH_3OH	CO	H_2O	C_2H_6	CH_3OCH_3
																combinations⇒	18	2	4	1	3	2	2							
																origin ⇓	a_1	a_3	a_4	a_5	a_6	a_8	a_9	a_{10}	a_{11}	a_{12}	a_{13}	a_{14}	a_{15}	a_{16}
	1														$\rightarrow m_2$	—	−1	1							−1					
		1													$\rightarrow m_3$	—		−1	1				1							
			1												$\rightarrow m_4$	—	1	−1		1				−1						
				1											$\rightarrow m_5$	—			1	−1	1									
						1									$\rightarrow m_7$	—	1				−1			−1		1				
							1								$\rightarrow m_8$	—	1		−1					−1				1		
								1							$\rightarrow m_9$	—	−2												1	
									1						$\rightarrow m_{10}$	—	−1		−1							1				
										1					$\rightarrow m_{11}$	—	−1				−1									1
											1				$\rightarrow m_{12}$	—	−1					1	−1	1						
													1		$\rightarrow m_{14}$	—					−1	1	−1			1				
														1	$\rightarrow m_{15}$	—	−1					−1		1			1			

Figure 10. Setup for the application of the algorithm on the methanol mechanism, after the elimination of a_2 (HO_2) and a_7 (H). In this and subsequent figures, rows (mechanisms) and columns (intermediates) that are no longer active are immediately removed rather than marked the way they were in the ammonia example (Figures 3 to 7). Thus, the rows of m_1, m_6, and m_{13}, and the columns of a_2 and a_7 have been removed. The terminal species, not shown in Figure 9, have been included in the setup. The intermediates a_5 (CH_3O_2H) and a_8 (CHO) will be eliminated in parallel, because they entail disjoint sets of mechanisms.

s_1 →	s_2 →	s_3 →	s_4 →	s_5 →	s_6 →	s_7 →	s_8 →	s_9 →	s_{10} →	s_{11} →	s_{12} →	s_{13} →	s_{14} →	s_{15} →	mechanism ⇓	species ⇒	CH_3	CH_3O_2	OH	CH_3O	CH_2O	CH_4	O_2	CH_3OH	CO	H_2O	C_2H_6	CH_3OCH_3
																combinations⇒	18	2	4	3	2							
																origin ⇓	a_1	a_3	a_4	a_6	a_9	a_{10}	a_{11}	a_{12}	a_{13}	a_{14}	a_{15}	a_{16}
	1														$\rightarrow m_2$	—	−1	1					−1					
		1													$\rightarrow m_3$	—		−1	1		1								
						1									$\rightarrow m_7$	—	1			−1		−1		1					
							1								$\rightarrow m_8$	—	1		−1			−1				1			
								1							$\rightarrow m_9$	—	−2										1		
									1						$\rightarrow m_{10}$	—	−1		−1					1					
										1					$\rightarrow m_{11}$	—	−1			−1									1
			1	1											$\rightarrow m_{16}$	a_5: m_4+m_5	1	−1	1	1		−1							
											1			1	$\rightarrow m_{17}$	a_8: $m_{12}+m_{15}$	−2				−1	2			1				
													1	1	$\rightarrow m_{18}$	a_8: $m_{14}+m_{15}$	−1			−1	−1	1		1	1				

Figure 11. Setup for the application of the algorithm on the methanol mechanism, after the elimination of a_5 (CH_3O_2H) and a_8 (CHO). Partial mechanisms and intermediates that are not active have been removed. Next, the intermediate a_9 (CH_2O) will be eliminated. The overlap of the mechanisms is such that it is not possible to remove two intermediates in parallel.

s_1 →	s_2 →	s_3 →	s_4 →	s_5 →	s_6 →	s_7 →	s_8 →	s_9 →	s_{10} →	s_{11} →	s_{12} →	s_{13} →	s_{14} →	s_{15} →	mechanism ⇓	species ⇒	CH_3	CH_3O_2	OH	CH_3O	CH_4	O_2	CH_3OH	CO	H_2O	C_2H_6	CH_3OCH_3
																combinations⇒	18	3	6	3							
																origin ⇓	a_1	a_3	a_4	a_6	a_{10}	a_{11}	a_{12}	a_{13}	a_{14}	a_{15}	a_{16}
	1														$\rightarrow m_2$	—	−1	1				−1					
						1									$\rightarrow m_7$	—	1			−1	−1		1				
							1								$\rightarrow m_8$	—	1		−1		−1				1		
								1							$\rightarrow m_9$	—	−2									1	
									1						$\rightarrow m_{10}$	—	−1		−1				1				
										1					$\rightarrow m_{11}$	—	−1			−1							1
			1	1											$\rightarrow m_{16}$	a_5: m_4+m_5	1	−1	1	1	−1						
		1									1			1	$\rightarrow m_{19}$	a_9: m_3+m_{17}	−2	−1	1		2			1			
		1											1	1	$\rightarrow m_{20}$	a_9: m_3+m_{18}	−1	−1	1	−1	1		1	1			

Figure 12. State of the algorithm, for the methanol mechanism, after the elimination of a_9 (CH_2O). Partial mechanisms and intermediates that are not active have been removed. Next, the intermediate a_3 (CH_3O_2) will be eliminated.

s_1 →	s_2 →	s_3 →	s_4 →	s_5 →	s_6 →	s_7 →	s_8 →	s_9 →	s_{10} →	s_{11} →	s_{12} →	s_{13} →	s_{14} →	s_{15} →	mechanism ⇓	species ⇒	CH_3	OH	CH_3O	CH_4	O_2	CH_3OH	CO	H_2O	C_2H_6	CH_3OCH_3
																combinations⇒	10	6	3							
																origin ⇓	a_1	a_4	a_6	a_{10}	a_{11}	a_{12}	a_{13}	a_{14}	a_{15}	a_{16}
						1									$\rightarrow m_7$	—	1		−1	−1		1				
							1								$\rightarrow m_8$	—	1	−1		−1				1		
								1							$\rightarrow m_9$	—	−2								1	
									1						$\rightarrow m_{10}$	—	−1	−1				1				
										1					$\rightarrow m_{11}$	—	−1		−1							1
	1		1	1											$\rightarrow m_{21}$	a_3: m_2+m_{16}		1	1	−1	−1					
	1	1									1			1	$\rightarrow m_{22}$	a_3: m_2+m_{19}	−3	1		2	−1		1			
	1	1											1	1	$\rightarrow m_{23}$	a_3: m_2+m_{20}	−2	1	−1	1	−1	1	1			

Figure 13. State of the algorithm, for the methanol mechanism, after the elimination of a_3 (CH_3O_2). Next, the intermediate a_6 (CH_3O) will be eliminated.

s_1	s_2	s_3	s_4	s_5	s_6	s_7	s_8	s_9	s_{10}	s_{11}	s_{12}	s_{13}	s_{14}	s_{15}	mechanism ⇓	species ⇒	CH_3	OH	CH_4	O_2	CH_3OH	CO	H_2O	C_2H_6	CH_3OCH_3
																combinations⇒	10	8							
→	→	→	→	→	→	→	→	→	→	→	→	→	→	→		origin ⇓	a_1	a_4	a_{10}	a_{11}	a_{12}	a_{13}	a_{14}	a_{15}	a_{16}
							1								$\rightarrow m_8$	—	1	−1	−1				1		
								1							$\rightarrow m_9$	—	−2							1	
									1						$\rightarrow m_{10}$	—	−1	−1			1				
	1	1									1			1	$\rightarrow m_{22}$	a_3: m_2+m_{19}	−3	1	2	−1		1			
	1		1	1		1									$\rightarrow m_{24}$	a_6: $m_{21}+m_7$	1	1	−2	−1	1				
	1		1	1						1					$\rightarrow m_{25}$	a_6: $m_{21}+m_{11}$	−1	1	−1	−1					1
	2	1	1	1									1	1	$\rightarrow m_{26}$	a_6: $m_{21}+m_{23}$	−2	2		−2	1	1			

Figure 14. Setup of the methanol mechanism, after the elimination of a_6 (CH_3O). The intermediate a_4 (OH) will be eliminated next.

s_1	s_2	s_3	s_4	s_5	s_6	s_7	s_8	s_9	s_{10}	s_{11}	s_{12}	s_{13}	s_{14}	s_{15}	mechanism ⇓	species ⇒	CH_3	CH_4	O_2	CH_3OH	CO	H_2O	C_2H_6	CH_3OCH_3
																combinations ⇒	5							
→	→	→	→	→	→	→	→	→	→	→	→	→	→	→		origin ⇓	a_1	a_{10}	a_{11}	a_{12}	a_{13}	a_{14}	a_{15}	a_{16}
								1							$\rightarrow m_9$	—	−2						1	
	1	1					1				1			1	$\rightarrow m_{27}$	a_4: m_8+m_{22}	−2	1	−1		1	1		
	1		1	1		1	1								$\rightarrow m_{28}$	a_4: m_8+m_{24}	2	−3	−1	1		1		
	1		1	1			1			1					$\rightarrow m_{29}$	a_4: m_8+m_{25}		−2	−1			1		1
	2	1	1	1			2						1	1	$\rightarrow m_{30}$	a_4: $2m_8+m_{26}$		−2	−2	1	1	2		
	1	1							1		1			1	$\rightarrow m_{31}$	a_4: $m_{10}+m_{22}$	−4	2	−1	1	1			
	1		1	1		1			1						$\rightarrow m_{32}$	a_4: $m_{10}+m_{24}$		−2	−1	2				
	1		1	1					1	1					$\rightarrow m_{33}$	a_4: $m_{10}+m_{25}$	−2	−1	−1	1				1
	2	1	1	1					2				1	1	$\rightarrow m_{34}$	a_4: $2m_{10}+m_{26}$	−4		−2	3	1			

Figure 15. Setup of the methanol mechanism, after the elimination of a_4 (OH). Elimination of the only remaining intermediate a_1 (CH_3), yields the final results of Figure 16.

s_1 →	s_2 →	s_3 →	s_4 →	s_5 →	s_6 →	s_7 →	s_8 →	s_9 →	s_{10} →	s_{11} →	s_{12} →	s_{13} →	s_{14} →	s_{15} →	⇓ mechanism	species ⇒ origin ⇓	CH_4 a_{10}	O_2 a_{11}	CH_3OH a_{12}	CO a_{13}	H_2O a_{14}	C_2H_6 a_{15}	CH_3OCH_3 a_{16}
	1		1	1			1			1					$\rightarrow m_{29}$	a_4: m_8+m_{25}	−2	−1			1		1
	2	1	1	1			2						1	1	$\rightarrow m_{30}$	a_4: $2m_8+m_{26}$	−2	−2	1	1	2		
	1		1	1		1			1						$\rightarrow m_{32}$	a_4: $m_{10}+m_{24}$	−2	−1	2				
	1		1	1		1	1	1							$\rightarrow m_{35}$	a_1: $m_{28}+m_9$	−3	−1	1		1	1	
	2	1	1	1		1	2				1			1	$\rightarrow m_{36}$	a_1: $m_{28}+m_{27}$	−2	−2	1	1	2		
	3	1	2	2		2	2		1		1			1	$\rightarrow m_{37}$	a_1: $2m_{28}+m_{31}$	−4	−3	3	1	2		
	2		2	2		1	1		1	1					$\rightarrow m_{38}$	a_1: $m_{28}+m_{33}$	−4	−2	2		1		1
	4	1	3	3		2	2		2				1	1	$\rightarrow m_{39}$	a_1: $2m_{28}+m_{34}$	−6	−4	5	1	2		

Figure 16. Final results for the methanol mechanism, after the elimination of a_1 (CH_3).

s_1	s_2	s_3	s_4	s_5	s_6	s_7	s_8	s_9	s_{10}	s_{11}	s_{12}	s_{13}	s_{14}	s_{15}	mechanism	direct?	mechanism in Table X of Happel *et al.* (1990)
	1		1	1			1			1					m_{29}	√	not found, because s_{11} was omitted
	2	1	1	1			2						1	1	m_{30}	√	m_2/r_0; m_3/r_0 (lines 5 and 7)
	1		1	1		1			1						m_{32}	√	m_1/d_4; m_3/d_4; m_4/d_4; m_6/d_4 (lines 2,6,8,10)
	1		1	1		1	1	1							m_{35}	√	not found, because s_9 was omitted
	2	1	1	1		1	2				1			1	m_{36}	√	m_4/r_0 (line 7)
	3	1	2	2		2	2		1		1			1	m_{37}	$=m_{36}+m_{32}$	—
	2		2	2		1	1		1	1					m_{38}	$=m_{29}+m_{32}$	—
	4	1	3	3		2	2		2				1	1	m_{39}	$=m_{30}+2m_{32}$	—

Figure 17. Analysis of the results for the methanol mechanism. Some of the final mechanisms are not direct, and they can be formed from the direct ones as shown. The last column shows the correspondence of the direct mechanisms constructed here to those of Happel *et al.* (1990).

application of the algorithm, shown in Figures 1.9 to 1.16, will be explained below, emphasizing in particular the differences arising from the directionality of the steps.

The initial arrangement for the application of the algorithm on this example is given in Figure 1.9, which shows the mechanisms m_1 to m_{15}, each using only the step with the same index. Terminal species have been omitted in this figure but will be included in Figure 1.10. The arrows below each step and to the left of each mechanism serve as indicators of directionality. In the previous example on ammonia (Figures 1.2 to 1.7), all the arrows are bidirectional ($\leftrightarrows$), since all the steps – hence all mechanisms as well – can have a net rate in either the forward or the reverse direction. For the methanol example, in Figure 1.9, all steps (and all mechanisms) must have a net forward rate; thus, all the arrows are unidirectional ($\rightarrow$). For other mechanisms, some steps may be unidirectional and others bidirectional; since the algorithm does depend on the directionality of steps, it is important to maintain these labels.

In the formation of combinations of partial mechanisms to eliminate an intermediate species, it is no longer possible to take any combination of two mechanisms whose net reactions involve the species. If, for example, we attempt to combine m_1 and m_{13} to eliminate a_2 (HO_2), then the combination expression given in the previous example:

$\beta_{b2}\, m_k - \beta_{k2}\, m_b$

would lead to either

$m_1 - m_{13}$ for k=1 and b=13

or, reversing the assignments of k and b, to:

$m_{13} - m_1$ for k=13 and b=1

In either case, we would be violating the directionality of one of the mechanisms m_1 and m_{13}. This happens because a_2 participates as a net product in both mechanisms; it is simply not possible to eliminate a_2 if we insist on using both mechanisms in the forward direction.

Thus, the rule for forming combinations must be modified: For unidirectional steps and mechanisms, a legitimate combination that eliminates an intermediate must include one mechanism in which the species is a net product and one mechanism in which the species is a net reactant. The number of combinations given in Figure 1.9 for each intermediate species reflects this requirement. The number of combinations is equal to yz, where y is the number of mechanisms for which the species in question is a net reactant and z the number of mechanisms for which the species is a net product. In other words, to obtain the number of combinations for each column corresponding to an intermediate we simply multiply the number of negative entries and the number of positive entries. We will later see that the procedure is somewhat more complicated for chemical systems that include both unidirectional and bidirectional steps.

As in the previous example, the intermediates with the smallest numbers of combinations are chosen for elimination. The species a_2 (HO_2) and a_7 (H) give rise to zero combinations, because they occur only as products – and never as net reactants. Thus, they are eliminated first, leading to the arrangement of Figure 1.10. Because of the size of this example, mechanisms that have been abolished and intermediates that have been eliminated are immediately removed from the figures; thus, the rows of m_1, m_6, and m_{13}, and the columns of a_2 and a_7 are not present in Figure 1.10 and subsequent figures. The number of combinations has been recalculated for each species. The intermediates a_5 (CH_3O_2H) and a_8 (CHO), each of which gives rise to only 2 combinations, are next eliminated in parallel, since their sets of mechanisms (m_4 and m_5 for a_5; m_{12}, m_{14}, and m_{15} for a_8) are disjoint. This results in the arrangement of Figure 1.11.

Next, elimination of the intermediate a_9 (CH_2O) which also involves 2 combination mechanisms leads to Figure 1.12. Similarly, elimination of the intermediate a_3 (CH_3O_2) leads to Figure 1.13, elimination of the intermediate a_6 (CH_3O) yields Figure 1.14, and elimination of a_4 (OH) Figure 1.15.

The final results are then obtained in Figure 1.16, after the elimination of the only remaining intermediate, a_1 (CH_3). Note that, throughout Figures 1.10 to 1.16, all combination coefficients in the origin column are positive, because reversal of the direction of a mechanism is not permitted.

The last three mechanisms that were produced (m_{37}, m_{38}, and m_{39}) are not direct; they can be formed from the direct ones as shown in Figure 1.17. The last column of Figure 1.17 shows the correspondence

of the direct mechanisms to those of Happel *et al.* (1990). It is important to note in the last column of Figure 1.17 that the same mechanism can be constructed many times in the procedure of Happel and Sellers (1983, p.290) – as well as the simple algorithm presented here. The issue of redundancies in the set of mechanisms is discussed further and addressed through a modified algorithm in Part 2. It should also be noted that, in Table X of Happel *et al.* (1990), the mechanisms m_1/r_0 (line 1), m_2/d_4 (line 4), m_5/r_0 (line 9),, m_5/d_4 (line 10), and m_6/r_0 (line 11) are all infeasible because they use either s_1 or s_5 in the wrong direction. The algorithm presented in this report never constructs mechanisms that violate the directionality of unidirectional steps.

With respect to the mechanisms m_{29} and m_{35}, which lead to byproducts, we observe (Figures 1.16 and 1.17) that m_{29} can be thought of as unrelated to the synthesis of methanol, but omission of m_{35} may be misleading: This mechanism leads to simultaneous production of methanol and ethane, in stoichiometric proportions. Thus, the mechanism m_{35} should properly be viewed as one of the mechanisms that lead to methanol – with the drawback, of course, that it also leads to an equal number of moles of an undesired byproduct.

DISCUSSION

In a chemical system that includes a set of species and a set of possible reaction steps, some species are usually considered as reaction intermediates, present only in small amounts during the course of the reaction. Other species are terminal species, and comprise the raw materials, products, and by-products of the system; these species can be consumed or produced in significant amounts. An *overall* mechanism accomplishes an *overall* reaction whose net stoichiometry involves only terminal species. Thus, overall reactions are responsible for the conversion of raw materials to final products and by-products; the buildup of the necessary small concentration of an intermediate, such as may be necessary in the initiation of a reaction, is not accomplished by overall mechanisms.

Direct mechanisms are a particular class of overall mechanisms; they are the shortest possible overall mechanisms, in the sense that if one reaction step is removed from a direct mechanism then no overall mechanism can be formed by the remaining steps. A mechanism which is indirect can always be decomposed into two or more direct mechanisms, each of which involves fewer steps than the original mechanism. Happel and Sellers (1982, 1983, 1989), Happel (1986), Sellers (1984, 1989), and Happel *et al.* (1990) introduced these concepts and considerations, established that the identification of direct reaction mechanisms for a given a set of reaction steps provides important insights in the behavior of chemical reaction systems, and presented a method for their systematic construction. That method of Happel and Sellers is based on identifying first a set of linearly independent reactions, with one mechanism for each reaction, and a set of linearly independent cyclic mechanisms, assuming all steps are reversible. The procedure then forms direct mechanisms as combinations of cyclic and non-cyclic mechanisms. In the end, mechanisms which utilize a step in a prohibited direction are rejected.

This report introduces an alternative approach which is based on starting with each reaction step considered as a partial mechanism. Then, one intermediate after another are examined, and the set of partial mechanisms is modified so that the intermediate does not appear in the net stoichiometry; the modification of mechanisms is carried out in a way that preserves the correct direction of irreversible reaction steps. By processing all intermediates in this way, a set of overall mechanisms is constructed.

The proposed algorithm is based on a method for the synthesis of biochemical pathways from bioreactions (Mavrovouniotis, 1989, Mavrovouniotis *et al.*, 1990): Mechanism steps correspond to individual bioreactions (usually catalyzed by enzymes), overall reactions correspond to net biotransformations, and reaction mechanisms correspond to biochemical pathways. There are, however, significant differences in the formulation of the problem: The classification of species as either intermediates or terminal species, as introduced by Happel and Sellers (1982, 1983), states whether a species may appear in the net stoichiometry but does not distinguish between appearance of the species as a reactant and appearance as a product; such distinctions are in fact not very meaningful in the procedure of Happel and Sellers since steps and mechanisms are generally considered reversible. The classification used by Mavrovouniotis *et al.* (1990) is more general; it permits separate specifications on each species as net reactant or net product. However, the classification of Happel and Sellers (1982, 1983) is clearer and more natural both chemically and mathematically, and has been adopted throughout the work presented here. Differences between the biochemical-pathway procedure and the reaction-mechanism procedure also exist in the mathematical details (specifically, the use of matrices in Figures 1.2-1.7 and 1.9-1.17). Finally, the way in which

reversible steps are treated is substantially different in the two domains. Most bioreactions are irreversible, and the reversible bioreactions were treated by Mavrovouniotis *et al.* (1990) as an exception: For each reversible bioreaction, the forward and the reverse direction were viewed as separate reactions. The approach introduced here, on the other hand, makes reversibility considerations an integral part of the procedure, through directionality labels which are taken into account in the construction of mechanisms.

The proposed mechanism-construction approach was presented, in this report, in the context of two examples, which have also been treated by Happel *et al.* (1990): An example involving reversible steps for the synthesis of ammonia, and an example with irreversible steps for the synthesis of methanol. In the methanol example, Happel *et al.* (1990) ignored the steps that lead to ethane and dimethylether (which are byproducts), assuming that they are irrelevant for methanol mechanisms. In the treatment of the methanol example in the present report, however, it was shown that there is one mechanism which involves simultaneous production of ethane and methanol (m_{35} Figure 1.16 and 17); thus, the omission of byproducts is a risky simplification.

It should be clear from the treatment of these examples that the two procedures operate in a totally different way, although many elements of the theory of direct mechanisms conceived by Happel and Sellers (1982, 1983), Happel (1986), and Sellers (1984, 1989) are important for the approach presented here. One should be careful to distinguish between a mathematical theory defining direct mechanisms and establishing their significance and a procedure for actually constructing direct mechanisms. It is in the latter where the two approaches differ drastically, as is further shown in Part 2 where the informal description of the algorithm given in this part is further formalized and the differences between the two methods are illustrated through a number of hypothetical chemical systems.

One particularly important difference between the procedure of Happel and Sellers and the approach presented here is the way in which they treat steps with fixed direction; such steps can be either thermodynamically irreversible or otherwise known or restricted to occur with a net forward rate. Happel and Sellers assume that all steps are reversible and construct mechanisms which can use any step in any direction; in the end, mechanisms violating the direction of any step are rejected. The approach presented here maintains and enforces requirements on the directionality of steps during the whole mechanism-construction process; complete or partial mechanisms that use an irreversible step in the wrong direction are never constructed. This difference is apparent in the second example (methanol) presented in this report. The inclusion of directionality considerations appears particularly cumbersome in the treatment of methanol mechanisms by Happel *et al.* (1990, p. 1061): The direct submechanisms for one of the basis reactions violate directionality restrictions; thus, an alternative reaction is manually constructed which gives acceptable direct submechanisms.

An apparent disadvantage of the approach suggested here is that the final set of mechanisms can include duplicate mechanisms and even indirect ones. This difficulty is addressed in Part 2 where procedures for eliminating such redundant mechanisms in the end, or even preventing their construction, are presented. It should be noted that in the procedure of Happel and Sellers similar steps must be taken to avoid duplicate mechanisms (Happel and Sellers, 1983, p. 290-291).

Finally, the present approach appears to deviate from that of Happel and Sellers (1982, 1983, 1989), Happel (1986), Sellers (1984, 1989), and Happel *et al.* (1990) in the definition of direct mechanisms. Specifically, Happel and Sellers consider a mechanism direct if it is minimal with respect to the reaction it accomplishes, i.e., if removing any step would no longer permit construction of a mechanism that accomplishes *the same reaction.* This report, on the other hand, defines direct mechanisms regardless of the reaction which they accomplish: A mechanism is direct if removal of any step would prohibit construction of any other *overall* mechanism. This difference cannot be analyzed in the context of the examples presented here, because it only arises when multiple non-null overall reactions are possible. Each of the two examples used here entails only one overall reaction. Examples in which this difference has consequences are considered in Part 2.

PART 2. FORMALIZATION AND ANALYSIS OF THE SYNTHESIS ALGORITHM

INTRODUCTION

A chemical reaction system includes a set of reaction steps and a set of species; each of the latter is classified as either a reaction intermediate, which cannot have large *net* consumption or production, or a terminal species, which can be consumed or produced in significant amounts. Reaction intermediates can be either unstable short-lived species, such as free radicals, or species which are generally stable but under the conditions of the chemical system are sufficiently reactive to be observed only in very small amounts. An *overall* reaction has only terminal species in its net stoichiometry. An *overall* mechanism is a linear combination of reaction steps which accomplishes an overall reaction. A mechanism is considered to advance at one particular reaction rate; the linear combination coefficients of the steps participating in the mechanism then dictate the rates of the steps.

Direct mechanisms (Milner, 1964) are the shortest possible overall mechanisms; indirect mechanisms can always be decomposed into direct mechanisms. Direct mechanisms have been extensively analyzed and used by Happel and Sellers (1982, 1983, 1989), Happel (1986), Sellers (1984, 1989), and Happel *et al.* (1990), and defined in slightly different form in Part 1 of this report. A sound justification of the chemical significance (Happel and Sellers, 1982, 1983, 1989, Happel, 1986, and Happel *et al.*, 1990) as well an elegant mathematical foundation of direct mechanisms (Sellers, 1984, 1989) have been provided by Happel and Sellers, along with a specific method for the construction of direct mechanisms. This method initially assumes that all steps are reversible and identifies one mechanism for each reaction within a basis set of linearly independent reactions. It also identifies a basis set of linearly independent cyclic mechanisms. Then, it forms direct mechanisms as combinations of these linearly independent mechanisms; finally, it rejects those mechanisms which include an irreversible step in the wrong direction.

An alternative approach, which has been introduced Part 1, starts with each reaction step considered as a partial mechanism and then transforms this set of partial mechanisms so that the intermediate species are eliminated from the net stoichiometry. Elimination of all intermediates yields in the end a set of overall mechanisms that includes all direct mechanisms and preserves the correct direction of irreversible reaction steps. The approach has been presented only informally, in the context of two simple examples involving chemical systems for the synthesis of ammonia and methanol.

One important feature of the approach is that it prohibits irreversible steps from being used in the wrong direction during the whole mechanism-construction process, i.e., even in incomplete, partial mechanisms. An apparent disadvantage of the approach is that the final set of mechanisms can include duplicate mechanisms (which can also be the case with the procedure of Happel and Sellers, 1983, p. 290) and even indirect mechanisms, which must be weeded out in the end. In addition to its mode of operation, the approach also deviates from that of Happel and Sellers (1982, 1983, 1989), Happel (1986), and Sellers (1984, 1989) in the definition of direct mechanisms.

The purpose of Part 2 of this report is to provide a more formal description of the basic algorithm for construction of direct mechanisms, as well as formal procedures for eliminating redundant (duplicate or indirect) mechanisms either at the very end or throughout the course of the algorithm. Some important issues, and in particular the effect of the somewhat different definition of direct mechanisms and the presence of irreversible reaction steps are examined within model studies involving hypothetical chemical systems. The proposed new procedure operates quite differently from that of Happel and Sellers (1982, 1983, 1989) and Happel (1986) and the formalization of the algorithm presented in this report, along with certain model studies, should establish this fact clearly. The underlying mathematical framework of direct mechanisms, elegantly set forth by Happel and Sellers (1982, 1983, 1989) and Sellers (1984, 1989) is important to both algorithms, but from the procedural viewpoint, i.e., in terms of the actual transformations and computations entailed by the two algorithms, the two approaches are quite distinct.

DESCRIPTION OF BASIC ALGORITHM

We now proceed to a formal description of the general algorithm whose fundamental approach was demonstrated through examples in Part 1. One aspect of the algorithm which was overly simplified in those examples is the selection of combinations of mechanisms to eliminate an intermediate. The first

example given in Part 1 on ammonia mechanisms contained only bidirectional (reversible) steps; the second example on methanol mechanisms contained only unidirectional (irreversible) steps. The formal description of the algorithm that is presented here will address, among other issues, the formation of combinations for the general case, i.e., when some steps are bidirectional and some unidirectional.

The description of the algorithm will follow the notation set forth by Happel and Sellers (1982, 1983, 1989), in the form it which it was presented in Part 1, including considerations related to the directionality of steps and mechanisms. Thus, we are considering a chemical system of I intermediate species, denoted as a_1 to a_I, T terminal species, denoted as a_{I+1} to a_A (where A=I+T), and S mechanism steps, denoted as s_1 to s_S. The stoichiometric coefficient of species a_j in step s_i is α_{ij}, and the transformation accomplished by step s_i is $r_i=R(s_i)=\sum_{j=1}^{A}\alpha_{ij}a_j$. Each step s_i labelled as either unidirectional (denoted as $\rightarrow s_i$) or bidirectional (denoted as $\leftrightarrows s_i$). If a reversible step is known to have a net forward rate under the conditions of interest, then it is treated as irreversible (unidirectional) in the application of the algorithm.. One may note that steps which proceed solely in the reverse direction are not permitted here; they can be trivially converted into forward unidirectional steps by changing the signs of the corresponding stoichiometric coefficients α_{ij}.

A mechanism (or mechanistic vector) m_k is a linear combination of steps, $m_k=\sum_{i=1}^{S}\sigma_{ki}s_i$, but with an additional restriction on σ_{ki}, to prohibit the use of unidirectional steps in the wrong direction:

$$\rightarrow s_i \;\Rightarrow\; \sigma_{ki}\geq 0$$

A mechanism may also have a directionality of its own, determined from the directionality of its steps. Specifically, if a mechanism m_k contains at least one unidirectional step s_i with a non-zero coefficient, then the mechanism is unidirectional; otherwise, the mechanism is bidirectional.

$$(\exists s_i:\ \rightarrow s_i \wedge \sigma_{ki}>0) \Leftrightarrow \rightarrow m_k$$

$$(\forall s_i \ni \rightarrow s_i:\ \sigma_{ki}=0) \Leftrightarrow \leftrightarrows m_k$$

The reaction vector r_k of the net transformation accomplished by mechanism m_k is $r_k=R(m_k)=\sum_{i=1}^{S}\sigma_{ki}R(s_i)=\sum_{j=1}^{A}\beta_{kj}a_j$, where $\beta_{kj}=\sum_{i=1}^{S}\sigma_{ki}\alpha_{ij}$. One may note that r_g for $g\leq S$ is ambiguous: It may be either the reaction of the step s_g or the reaction of the mechanism m_g. This ambiguity is harmless here, because each of the first S mechanisms in the algorithm will contain only the corresponding step with the same index.

A mechanism m_k is an overall mechanism iff does not involve intermediates in its net reaction stoichiometry, i.e., has $\beta_{kj}=0$ for j=1 to I. An overall mechanism m_k is considered here direct iff no overall mechanism exists using only a subset of the steps that appear with non-zero coefficients in m_k. In contrast to Happel and Sellers (1982, 1983, 1989), Happel (1986), Sellers (1984, 1989), and Happel *et al.* (1990), we do not define direct mechanisms in terms of the particular overall reaction which they accomplish. They define a mechanism m_k as *direct for a particular reaction r*, when no other mechanism *for r* involves a subset of the steps of m_k. The definition used in this report, on the other hand, makes no particular reference to the reaction accomplished by the mechanism, other than the requirement that only *overall* mechanisms be considered. The difference can be traced to the operation of the two algorithms. Happel and Sellers first derive a basis set of reactions and then construct mechanisms for particular reactions. The algorithm presented here proceeds directly to the construction of mechanisms (the analysis of reactions can be carried out at the end). Some issues related to the two options in the definition of

s_1	s_2	.....	s_S		number of combinations ⇒	n_1	n_2	.....	n_I	–	.....	–
ε_1	ε_2	.....	ε_S	mechanism ⇓	origin ⇓	a_1	a_2	.....	a_I	a_{I+1}	.....	a_A
σ_{11}	σ_{12}	.....	σ_{1S}	ε_1 m_1	—	β_{11}	β_{12}	.....	β_{1I}	$\beta_{1(I+1)}$	.....	β_{1A}
σ_{21}	σ_{22}	.....	σ_{2S}	ε_2 m_2	—	β_{21}	β_{22}	.....	β_{2I}	$\beta_{2(I+1)}$	.....	β_{2A}
⋮	⋮	⋮	⋮	⋮ ⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
σ_{S1}	σ_{S2}	.....	σ_{SS}	ε_S m_S	—	β_{S1}	β_{S2}	.....	β_{SI}	$\beta_{S(I+1)}$	.....	β_{SA}

Figure 2.1 Initial setup for the application of the algorithm for the construction of mechanisms. The left portion of the table contains the coefficients σ_{ki} indicating the coefficient of step s_i in mechanism m_k; initially, each mechanism m_k contains only the corresponding step s_k, hence $\sigma_{ki}=1$ if $k=i$ while $\sigma_{ki}=0$ if $k\neq i$. Below each step symbol s_i there is the label ε_i which identifies the directionality of the step; this label is either $\rightarrow$, denoting a unidirectional step, or $\leftrightarrows$, denoting a bidirectional step. The label of mechanism m_k is initially identical to that of the corresponding step. The right portion of the table shows the reactions accomplished by the mechanisms, and groups intermediate species as a_1 to a_I, and terminal species as a_{I+1} to a_A. Since initially each mechanism consists of just the corresponding step the entries in the right portion of the table are just stoichiometries of steps, i.e., $\beta_{kj}=\alpha_{ki}$. The number of combination mechanisms that must be created to eliminate each species is listed above the species symbol. Numbers of combinations are not shown for terminal species, because they will not be eliminated. The column marked "origin" in the middle of the table is not essential; it is used for keeping track of how each new mechanism is constructed; one can list the intermediate that was eliminated and the mechanisms that were combined in the construction.

Note

A statement showing formation of a mechanism as combination of other mechanisms implies proper update of the β and σ matrices. Thus, the statement:

$m_c := \beta_{bJ}\, m_k - \beta_{kJ}\, m_b$

should be understood to imply also:

For each $a_j \in (N \cup N_T)$ do:

$\beta_{cj} := \beta_{bJ}\, \beta_{kj} - \beta_{kJ}\, \beta_{bj}$

For each i=1, ...,S do:

$\sigma_{ci} := \beta_{bJ}\, \sigma_{ki} - \beta_{kJ}\, \sigma_{bi}$

1. Initialization

$N := \{a_1, a_2, ..., a_I\}$
$M := \{m_1, m_2, ..., m_S\}$
For k=1,...,S do:
  For i=1,...,S do:
    If k=i then do:
      $\sigma_{ki} := 1$
    Else do:
      $\sigma_{ki} := 0$
  If $\rightarrow s_k$ then do:
    Assign directionality $\rightarrow m_k$
  Else do:
    Assign directionality $\leftrightarrows m_k$
  For j=1,...,A do:
    $\beta_{kj} := \alpha_{kj}$

2. Number of combinations

For each $a_j \in N$ do:
  $y_j := 0;\ z_j := 0;\ x_j := 0$
  For each $m_k \in M$ such that $\beta_{kj} \neq 0$ do:
    If $\leftrightarrows m_k$ then do:
      $x_j := x_j + 1$
    Else if $\beta_{kj} < 0$ then do:
      $y_j := y_j + 1$
    Else if $\beta_{kj} > 0$ then do:
      $z_j := z_j + 1$
  $n_j := x_j(x_j - 1)/2 + x_j z_j + x_j y_j + z_j y_j$

3. Elimination of intermediate

$n_{min} := +\infty$
For each $a_j \in N$ do:
  If $n_j < n_{min}$ then do:
    $n_{min} = n_j;\ J := j$
$Y_J := \emptyset;\ Z_J := \emptyset;\ X_J := \emptyset$
For each $m_k \in M$ such that $\beta_{kJ} \neq 0$ do:
  If $\leftrightarrows m_k$ then do:
    $X_J := X_J \cup \{m_k\}$
  Else if $\beta_{kJ} < 0$ then do:
    $Y_J := Y_J \cup \{m_k\}$
  Else if $\beta_{kJ} > 0$ then do:
    $Z_J := Z_J \cup \{m_k\}$
$M_J := \emptyset$
For each $m_b \in X_J$ do:
  For each $m_k \in X_J$ such that k<b do:
    $m_c := \beta_{bJ}\, m_k - \beta_{kJ}\, m_b$
    Assign directionality $\leftrightarrows m_c$
    $M_J := M_J \cup \{m_c\}$
  For each $m_k \in (Z_J \cup Y_J)$ do:
    If $\beta_{bJ} > 0$ then do:
      $m_c := \beta_{bJ}\, m_k - \beta_{kJ}\, m_b$
    Else do:
      $m_c := \beta_{kJ}\, m_b - \beta_{bJ}\, m_k$
    Assign directionality $\rightarrow m_c$
    $M_J := M_J \cup \{m_c\}$
For each $m_k \in Y_J$ do:
  For each $m_b \in Z_J$ do:
    $m_c := \beta_{bJ}\, m_k - \beta_{kJ}\, m_b$
    Assign directionality $\rightarrow m_c$
    $M_J := M_J \cup \{m_c\}$

4. Update of active sets

$M := (M - (X_J \cup Y_J \cup Z_J)) \cup M_J$
$N := N - \{a_J\}$

5. Check for termination

If $N \neq \emptyset$ then go back to phase 2
Else return M.

Figure 2.2. A succinct statement of the basic algorithm for the construction of reaction mechanisms from a set of mechanism steps. Appropriate indenting makes the scope of conditionals and iteration constructs clear. A detailed explanation of each phase and operation in the algorithm is provided in the text.

direct mechanisms are discussed in a subsequent section in this report, in the context of a hypothetical chemical system.

The algorithm operates on a set of mechanisms, M, a set of intermediate species, N, and the set of terminal species, N_T. During the operation of the algorithm, the mechanisms in the set M are not necessarily overall mechanisms, and they may be referred to as *partial* mechanisms; the set M is iteratively updated by the algorithm through both removal of mechanisms and addition of newly-constructed ones, aiming to yield, ultimately, a set of direct overall mechanisms. In addition to M, N, and N_T the state of the algorithm, updated in each iteration includes: The directionality labels of the mechanisms in M; the numbers of combinations, n_j, which are computed in the course of the algorithm for the intermediates in N; and the matrices of β_{kj} and σ_{ki} — for i=1,...,S and those indices k and j for which $m_k \in M$ and $a_j \in (N \cup N_T)$.

The application of the algorithm can be carried out using the layout of Figure 2.1, which arranges all the necessary information in a convenient format and explains how the entries are initialized; this is the same layout that was used in the examples of Part 1. A formal statement of the algorithm is shown in Figure 2.2, which makes the proper nesting of operations such as loops and conditionals clear. The description given below explains the individual operations involved in more detail. The algorithm is described here in its most basic form and without formal statements and proofs of its properties. Some refinements and variations of the algorithm are discussed in a subsequent section.

Algorithm 1 (Basic algorithm)

Phase 1. Initialization: Let N be the set of all the intermediate species:

$N:=\{a_1, a_2, ..., a_I\}$

We will assume that $N \neq \emptyset$; if $N=\emptyset$ then there are no intermediate species and each individual step is a direct overall mechanism.

Let M be the set of S mechanisms of length 1:

$M:=\{m_1, m_2, ..., m_S\}$

Their coefficients are:

$\sigma_{ki}=\delta_{ki}$, k=1 to S, i=1 to S

where δ_{ki} is the Kronecker delta ($\delta_{ki}=1$ if k=i; $\delta_{ki}=0$ if $k \neq i$). From the definition of mechanism directionality, it can be easily shown that each of these mechanisms has the same directionality as the corresponding step:

$\rightarrow s_k \Leftrightarrow \rightarrow m_k \quad (k=1, ..., S)$

$\leftrightarrows s_i \Leftrightarrow \leftrightarrows m_k \quad (k=1, ..., S)$

Also, from the definition of β_{kj}:

$\beta_{kj}=\alpha_{kj} \quad (k=1, ..., S)$

The setup of Figure 2.1 explains the initialization and shows a convenient way of arranging all the information.

Phase 2. Computation of the number of combinations for each intermediate: For each $a_j \in N$ do:

Let Y_j be the set of irreversible mechanisms in whose net reactions a_j appears as a reactant.

$Y_j=\{m_k / \ m_k \in M \wedge \rightarrow m_k \wedge \beta_{kj}<0\}$

Let Z_j be the set of irreversible mechanisms in whose net reactions a_j appears as a product.

$Z_j=\{m_k / \ m_k \in M \wedge \rightarrow m_k \wedge \beta_{kj}>0\}$

Let X_j be the set of reversible mechanisms in whose net reactions a_j participates with a non-zero coefficient.

$X_j=\{m_k / \ m_k \in M \wedge \leftrightarrows m_k \wedge \beta_{kj}\neq 0\}$

Let y_j, z_j, and x_j be the cardinalities of the sets:

$y_j = |Y_j|; \ z_j = |Z_j|; \ x_j = |X_j|$

Referring to Figure 2.1, one can obtain the numbers y_j, z_j, and x_j by scanning the column corresponding to a_j and noting the non-zero entries; each non-zero entry, is counted into one of the the three categories, depending on its sign and the directionality label corresponding to it (in the mechanism column). For example, a positive entry corresponding to a unidirectional mechanism is counted into z_j.

Thus, the species a_j participates in y_j irreversible mechanisms as a net reactant, z_j irreversible mechanisms as a net product, and x_j reversible mechanisms. The number of combinations n_j is computed as:

$n_j=x_j(x_j-1)/2+x_jz_j+x_jy_j+z_jy_j$

and can be placed in the appropriate location in the setup of Figure 2.1. The examples that were used in Part 1 showed two extreme cases of this formula: If all steps are reversible, then $z_j=y_j=0$, and there are $x_j(x_j-1)/2$ combinations (ammonia example); if all the steps are irreversible then $x_j=0$, and there are only z_jy_j combinations (methanol example).

Phase 3. Selection and elimination of an intermediate: Let a_J be the intermediate with the smallest number of combinations. In the case of a tie, a_J can be any of the intermediates which entail the (identical) minimum number of combinations; however, to ensure reproducibility of results, we will take the intermediate with the smallest index:

$J=\min(J')$ where $n_{J'}=\min_j(n_j)$

It will be shown later that only efficiency considerations dictate the choice of intermediate, and that the rule used above is by no means the only plausible one. In fact, the basic algorithm can produce all direct mechanisms even if the intermediates are processed in random order, albeit with more effort. Refinements of the algorithm, discussed in companion papers, make the choice even less important.

Let M_J be the set of new mechanisms which are constructed as follows:

- For each combination m_k, m_b, where $m_k \in X_J$, $m_b \in X_J$, and $k<b$, form a combination mechanism, m_c:

$m_c=\beta_{bJ} m_k - \beta_{kJ} m_b$

and, since it is constructed as a combination of bidirectional mechanisms, assign a bidirectional label to it:

$\leftrightarrows m_c$

The restriction $k<b$ is needed to exclude the combination of a mechanism with itself (i.e., $k=b$) and also to make each combination (rather than permutation) of two mechanisms from X_J appear only once. As long as each combination of two distinct mechanisms from X_J is used precisely once, it is not important to observe the restriction $k<b$; in effect, either the combination $\beta_{bJ} m_k - \beta_{kJ} m_b$ or the combination $\beta_{kJ} m_b - \beta_{bJ} m_k$ should be formed (but not both). A total of $x_J(x_J-1)/2$ such combinations exist.

- For each combination m_k, m_b, where $m_k \in Z_J$, $m_b \in X_J$, form a unidirectional combination-mechanism, m_c, as follows:

If $\beta_{bJ}>0$, form: $m_c=\beta_{bJ} m_k - \beta_{kJ} m_b$ and assign $\rightarrow m_c$

If $\beta_{bJ}<0$, form: $m_c=\beta_{kJ} m_b - \beta_{bJ} m_k$ and assign $\rightarrow m_c$

The mechanism m_c must be unidirectional because its formation includes the unidirectional mechanism m_k. The second of the two alternatives above is merely the reverse of the first. The two alternatives are used to ensure that the coefficient which multiplies m_k is positive; since m_k is a unidirectional mechanism, it is not

permitted to be used with a negative coefficient. A total of $x_J z_J$ such combination mechanisms exist in this group.

• Similarly, for each combination m_k, m_b, where $m_k \in Y_J$, $m_b \in X_J$, form a unidirectional mechanism, m_c, as follows:

If $\beta_{bJ}>0$, form: $m_c=\beta_{bJ}\, m_k - \beta_{kJ}\, m_b$ and assign $\rightarrow m_c$

If $\beta_{bJ}<0$, form: $m_c=\beta_{kJ}\, m_b - \beta_{bJ}\, m_k$ and assign $\rightarrow m_c$

The two alternatives again ensure that m_k, a unidirectional mechanism is used in the proper direction. A total of $x_J y_J$ such combination mechanisms exist in this group.

• Finally, for each combination m_k, m_b, where $m_k \in Y_J$, $m_b \in Z_J$, form the unidirectional mechanism:

$m_c=\beta_{bJ}\, m_k - \beta_{kJ}\, m_b$ and assign $\rightarrow m_c$

The definition of the set Z_J guarantees that if $m_b \in Z_J$ then a_J is a product of $R(m_b)$, i.e., $\beta_{bJ}>0$ always. Similarly, the definition of Y_J guarantees that if $m_k \in Y_J$ then a_J is a reactant of $R(m_k)$, i.e., $\beta_{kJ}<0$, hence $-\beta_{kJ}>0$. Consequently, the coefficients of the mechanisms m_k and m_b in the combination are both positive and thus preserve the directionality requirements of the constituent mechanisms. A total of $z_J y_J$ such combination mechanisms exist in this group.

• The total number of combinations for all of the above categories is n_J:

$n_J=x_J(x_J-1)/2+x_J z_J+x_J y_J+z_J y_J$

Thus, the formation of the combinations detailed above justifies the formula used for n_j in the previous phase of the algorithm.

• The formation of each combination is accomplished by carrying out the linear combination of the coefficients β_{kj} and σ_{ki} of the constituent mechanisms to obtain the respective coefficients of the combination mechanism. In effect, for a mechanism formed as $m_c=\beta_{bJ}\, m_k - \beta_{kJ}\, m_b$:

$\forall j \ni a_j \in (N \cup N_T)$: $\beta_{cj}=\beta_{bJ}\, \beta_{kj} - \beta_{kJ}\, \beta_{bj}$

$\forall i=1, ...,S$: $\sigma_{ci}=\beta_{bJ}\, \sigma_{ki} - \beta_{kJ}\, \sigma_{bi}$

Following the first equation, the coefficient for a_J in each of the new combination mechanisms m_c has the general form $\beta_{cJ} = \beta_{bJ}\beta_{kJ}-\beta_{kJ}\beta_{bJ}$; the resulting coefficient is zero, i.e., a_J has indeed been eliminated.

In reference to Figure 2.1, the formation of a new mechanism as a linear combination of existing ones simply means that a new row is inserted at the bottom of the setup, and (except for the columns "mechanism" and "origin") is computed as a linear combination of the rows of the mechanisms being combined.

The "origin" entry is merely for bookkeeping purposes and one can include in it the name of the intermediate being eliminated and a two-term expression denoting the linear combination of mechanisms that was used in the construction of the new mechanism. For example, "a_1: $4m_3$–m_9" would mean that the row in question was created as a linear combination of m_3 (with a coefficient of 4) and m_9 (with a coefficient of –1, which implies that m_9 is bidirectional), during the elimination of intermediate a_1. This format was used in Part 1 in examples (Figures 1.3-1.7, 1.11-1.16). The "mechanism" entry includes a symbol or index for the new mechanism, and an appropriate directionality label which is $\leftrightarrows$ if both mechanisms used in the combination had the label $\leftrightarrows$, but $\rightarrow$ if at least one constituent mechanism has the label $\rightarrow$.

Phase 4. Update of active sets: The set of new combination mechanisms M_J is added to M, and the set of constituent mechanisms that were used are removed from M. Hence, the new set of active mechanisms is equal to:

$M:=(M\cup M_J)-(X_J\cup Y_J\cup Z_J)$

The intermediate a_J is removed from N, i.e.,

$N:=N-\{a_J\}$

In the arrangement of Figure 2.1, in this phase one first removes (crosses out or marks as "inactive" in some other way) as all rows that have non-zero entries in the column corresponding to a_J and then removes the column of a_J itself.

Phase 5. Check for termination: If $N\neq\emptyset$, i.e., if in Figure 2.1 no active columns remain corresponding to intermediates, then we go back to phase 2 (computation of numbers of combinations). If $N=\emptyset$ then the elimination of intermediates has been completed, and the resulting set of mechanisms, M, contains all direct mechanisms of the system. However, M can also contain duplicate mechanisms and even indirect ones. This issue is addressed further in a subsequent section.

IDENTIFICATION AND PRUNING OF REDUNDANT MECHANISMS

The set of mechanisms, M, resulting from the application of the basic algorithm discussed above contains all direct mechanisms of the system. However, M can also contain duplicate mechanisms (as in the ammonia example) or even indirect ones (as in the methanol example). This section will show a simple procedure for eliminating redundant mechanisms.

The length of a mechanism has been defined as the number of steps which participate in the mechanism (with non-zero coefficients) and is denoted as $|T(m_k)|$. Consider the set M partitioned into classes based on mechanism length. This partitioning is useful because indirect mechanisms are longer than any of the direct mechanisms from which they can be constructed (since direct mechanisms are minimal). With u a positive integer, let L(u,M) represent the set of all u-length mechanisms in M:

$$L(u,M)=\{m_k: m_k\in M \wedge |T(m_k)|=u\}$$

Let u_{max} be the largest mechanism length:

$$u_{max}(M)=\max_{m_k\in M} |T(m_k)|$$

The algorithm shown below converts the final set of mechanisms M into a revised set M' which excludes redundant (duplicate or indirect) mechanisms. The classes L(u,M) are processed in succession.

Algorithm 2 (Batch elimination of redundant mechanisms)

Phases 1-5: Identical to Algorithm 1.

Phase 6. Elimination of duplicate or indirect mechanisms:

Compute $u_{max}(M)=\max_{m_k\in M} |T(m_k)|$.

Compute $L(u,M)=\{m_k: m_k\in M \wedge |T(m_k)|=u\}$ for $u=1,\ldots, u_{max}(M)$.

Initialize $M':=\emptyset$.

For each $u=1,2,\ldots, u_{max}(M)$, consider each $m_k\in L(u,M)$ in succession, carrying out the following procedure:

- Check whether any mechanism $m_b\in M'$ contains a subset of the steps of m_k; if so, disregard m_k. Otherwise, add m_k to M'. Mathematically, iff

$$\nexists m_b\in M' \ni \quad T(m_b)\subseteq T(m_k)$$

then set

$M':=M'\cup\{m_k\}$

This procedure appears to entail two iteration loops, one for u and one for $m_k\in L(u,M)$, only because we wish to emphasize the correct order in the processing of mechanisms: The mechanisms in M are considered from shortest to longest. It should be clear, however, that each mechanism in M is scanned precisely once. Thus, the computational burden associated with this cleanup phase is modest.

To carry out the cleanup in the setup of Figure 2.1, one ranks the rows (mechanisms), based on the number of non-zero entries in the left (σ_{ki}) portion of the setup. Then, one considers each row in succession, beginning with the one that has the fewest non-zero σ_{ki} entries; each row will be either approved or rejected as follows. One inspects other rows that have already been approved. If there is an approved row which has non-zero σ_{ki} entries only for steps (columns) for which the row at hand has a non-zero σ_{ki}, then the row at hand is rejected; otherwise it is approved.

One does not need to postpone this cleanup of redundant pathways until the end. Indeed, if a significant number of redundant pathways is formed early on, they may then be combined to each other and lead to combinatorial explosion. To safeguard against such an event, the procedure given above can be applied incrementally, with each new mechanism introduced. In other words, a mechanism can be tested as shown in the algorithm immediately after it is constructed and before it is introduced in M. In this incremental checking, the new mechanism should only be compared to those mechanisms in M that will not be deleted to eliminate the intermediate which is being processed. In the symbols and terminology used in Phase 4 of Algorithm 1, the mechanisms in M_J should be first constructed; then M should be set to $M-(X_J \cup Y_J \cup Z_J)$, i.e., mechanisms that include a_J should be deleted; finally, the mechanisms in M_J should be introduced into M, one by one, provided they satisfy the condition posed in Algorithm 2, i.e., m_k is introduced into M iff $\nexists m_b \in M \ni T(m_b) \subseteq T(m_k)$. For clarity, then, the modification suggested for Algorithm 1 is as follows:

Algorithm 3 (Incremental Elimination of Redundant Mechanisms)

Phases 1-3: Identical to those in Algorithm 1.

Phase 4. Update of active sets: Mechanisms that include a_J are deleted from M, by setting:

$M:=M-(X_J \cup Y_J \cup Z_J)$.

The mechanisms in M_J are ranked by length. For each mechanism m_k in M_J, from shortest to longest, the following procedure is carried out:

- If $\nexists m_b \in M \ni T(m_b) \subseteq T(m_k)$, then $M:=M \cup \{m_k\}$; otherwise, m_k is disregarded.

Finally, the intermediate a_J is removed from N, i.e.,

$N:=N-\{a_J\}$

Phase 5: Identical to that in Algorithm 1.

The algorithmic efficiency of checking whether there exists another mechanism m_b which makes m_k redundant, i.e., finding an m_b for which $T(m_b) \subseteq T(m_k)$, can be enhanced by placing the mechanisms of M' (or M if the checking is to be done incrementally) in a directed graph, with arcs representing set-subset relationships. Alternatively, a binary tree of depth S, in which mechanisms are classified at the first level based on whether they involve a_1, then at the second level based on whether they involve a_2, etc. This arrangement would allow m_k to be matched against each m_b in M (or M') simultaneously rather than sequentially. These are only some of the available computer-implementation options and their details will not be discussed in this report; they are the subject of a separate report currently in preparation.

Within the setup of Figure 2.1, Algorithm 3 simply means that a new row (mechanism that is created) will be either approved or immediately discarded as follows. If there already exists another row which has non-zero σ_{ki} entries only for steps (columns) for which the new row has a non-zero σ_{ki}, then the new row is discarded; otherwise it is approved. In accordance with the algorithm, of course, one must delete the rows that have a non-zero entry under a_J *before* considering the new rows.

In the examples of methanol and ammonia discussed in Part 1, only a small number of redundant mechanisms were produced in the first place. In our experience, this is generally the case. The ordering of the intermediates in the main algorithm (processing the ones that create fewer combinations first) plays an important role in limiting the construction of redundant mechanisms.

SOME ISSUES AND VARIATIONS IN THE ALGORITHM

This section addresses a number of issues and variations that can affect the operation of the algorithm. They are addressed in this separate section to avoid entanglement of the description of the basic algorithm.

Selection of intermediate a_J for elimination. When two or more intermediates entail the same number of combinations (n_J), the algorithm chooses to eliminate the intermediate with the lowest index. This is clearly a completely arbitrary choice; any other tiebreaker rule would be satisfactory. A sensible modification would be to consider, in addition to n_J, the number of mechanisms already in M that will be dropped when a_J is eliminated. Let w_J be this number; we have:

$$w_J = |X_J \cup Y_J \cup Z_J| = |X_J| + |Y_J| + |Z_J| = x_J + y_J + z_J$$

since X_J, Y_J, and Z_J are disjoint; within Figure 2.1, w_J is merely the number of non-zero β_{kJ} entries in the column of a_J. While it is advantageous to choose an a_J such that n_J is as small as possible (to have the fewest possible new mechanisms to construct), it is also advantageous to have a large w_J, to remove as many mechanisms as possible from the set M, as this may simplify the subsequent elimination of other intermediates. This rule may remove some ties but is not guaranteed to identify a_J uniquely; thus, the preference for the smallest J will be demoted but still maintained. Based on these arguments, one can modify any of the Algorithms 1 to 3 using:

Variation 1 (Secondary selection of intermediate a_J based on the number of existing mechanisms that will be eliminated)

- The state of the algorithm includes, for each $a_j \in N$ the number of mechanisms in M in which it occurs as a net reactant or product; this number, w_j, can be defined as $|\{m_k / m_k \in M \wedge \beta_{kj} \neq 0\}|$.
- In **Phase 2** the computation $n_j := x_j(x_j-1)/2 + x_j z_j + x_j y_j + z_j y_j$ is followed by the computation $w_j := x_j + y_j + z_j$ (for each $a_j \in N$).
- In **Phase 3** the rule for selecting the target intermediate a_J is replaced by:

  $J = \min(J'')$ for the J'' for which $w_{J''} = \max_{J'}(w_{J'})$ for those J' satisfying $n_{J'} = \min_j(n_j)$.

  This rule simply states that one first considers only all those intermediates that give the smallest n_j then prunes this set down to those that give the smallest w_j and then picks the one with the smallest index.

This variation is meaningful only if there are some unidirectional steps. If all steps are bidirectional then $n_j = x_j(x_j-1)/2$ and $w_j = x_j$. Hence, a tie in n_j implies a tie in x_j and w_j as well, rendering the above variation useless. As an example in which the variation applies, consider one intermediate with $x_1=0$, $y_1=2$, $z_1=2$ and another with $x_1=0$, $y_1=4$, $z_1=1$. We obtain $n_1=n_2=4$, i.e., the original selection rule yields a tie; on the other hand, $w_1=4$ while $w_2=5$, i.e., the variation would lead to selection of a_2 over a_1.

The reason primary consideration in Variation 1 is given to n_j and w_j is used only as a secondary criterion is that it is generally desirable to avoid the computationally costly construction of mechanisms. If one focuses exclusively on maintaining the cardinality of M as low as possible after each iteration, then one would consider the net change in |M|. This net change is equal to $n_J - w_J$ if a_J is to be eliminated.

Variation 2 (Selection of intermediate a_J exclusively from the net change in the number of active mechanisms)

- The state of the algorithm includes, w_j, as in variation 1.
- In **Phase 2** the computation of n_j is followed by the computation of w_j as in variation 1.
- In **Phase 3** the rule for selecting the target intermediate a_J is replaced by:

  $J = \min(J')$ among those J' satisfying $n_{J'} - w_{J'} = \min_j(n_j - w_j)$.

Variation 2, like Variation 1, is meaningful only if some steps are unidirectional. In the example considered above, $n_1=n_2=4$, $w_1=4$, and $w_2=5$, one would obtain $n_1-w_1=0$ and $n_2-w_2=-1$ and again select a_2 over a_1. Consider another example, in which $x_1=0$, $y_1=2$, $z_1=2$ while $x_1=0$, $y_1=5$, $z_1=1$; here, $n_1<n_2$, but $n_1-w_1>n_2-w_2$. Hence, Variation 2, unlike Variation 1, may even lead to a clear-cut reversal of the order of processing the intermediates (regardless of arbitrary indexing)

One could, of course, choose a_J without any regard to either n_J or w_J. This would not hinder the the algorithm from constructing all mechanisms, but it would have a detrimental effect on the efficiency of the algorithm. Under Algorithms 1 to 3 and Variations 1 and 2, intermediates which have $n_j=0$ are immediately discarded, and so are all steps that involve these intermediates. If other intermediates are

instead eliminated then mechanisms may be constructed only to be discarded later. For example, in Figure 1.9 in Part 1, intermediates a_2 and a_7 have $n_2=n_7=0$ and are processed first, leading to the removal of m_1, m_6, and m_{13}. If one chooses instead to process a_1 first, one will construct 24 new mechanisms, 6 of which will utilize m_1. These six mechanisms and any mechanisms subsequently derived from them, will ultimately be discarded when a_2 is considered. Poor selection of intermediates can also lead to many redundant mechanisms constructed only to be rejected by the procedures of Algorithms 2 and 3. Examination of the model studies with hypothetical mechanisms in a subsequent further establishes the importance of the selection rule.

Unification of the selection of combinations. In the basic algorithm, the sets X_j, Y_j, and Z_j are assembled and treated separately to ensure that all constructed mechanisms are directionally feasible and are assigned the correct directionality label; likewise, the cardinalities x_j, y_j, and z_j allow an accurate count of the burden that elimination of an intermediate a_j entails. The algorithm can be made significantly simpler, at a loss of some efficiency, by lumping X_j, Y_j, and Z_j into a single set simply defined as:

$$W_j=\{m_k /\ m_k\in M \wedge \beta_{kj}\neq 0\}=X_j\cup Y_j\cup Z_j$$

This set, along with its cardinality w_j, is used in Variations 1 and 2 only for the purpose of ranking the intermediates. Here, it will be used in the construction of combinations that eliminate a selected intermediate a_J as follows. For each pair of mechanisms $m_k\in W_J$ and $m_b\in W_J$, $m_k\neq m_b$, consider the candidate new mechanisms $\beta_{bJ}\, m_k-\beta_{kJ}\, m_b$ and $\beta_{kJ}\, m_b-\beta_{bJ}\, m_k$ (the second being merely the reverse of the first). If both are directionally infeasible then they are rejected. Otherwise, a directionally feasible one (either one, if they are both feasible) is accepted. The variation proposed here is actually equivalent to the basic algorithm if all steps are bidirectional.

In this procedure, directionality labels for mechanisms could be omitted altogether: Each new mechanism would then be accepted as directionally feasible based on directionality of steps alone. Note also that the ordering of the intermediates would have to be based solely on w_j, with priority given to the lowest value; this ordering may lead to inefficiencies when many steps are unidirectional.

This variation is very different from the basic algorithm and is best presented as a separate algorithm.

Algorithm 4 (Simplified Consideration of Directionality)

Phase 1. Initialization: As in Algorithm 1, let $N=\{a_1, a_2, ..., a_I\}$ and assume that $N\neq\emptyset$. Let M be the set of S mechanisms of length 1, $M=\{m_1, m_2, ..., m_S\}$, whose coefficients are $\sigma_{ki}=\delta_{ki}$ and $\beta_{kj}=\alpha_{kj}$. Directionalities are not considered here.

Phase 2. Computation of the number of occurrences for each intermediate: For each $a_j\in N$ let $W_j=\{m_k /\ m_k\in M \wedge \beta_{kj}\neq 0\}$ be the set of mechanisms in whose net reactions involve a_j, and let $w_j=|W_j|$.

Phase 3. Selection and elimination of an intermediate: Let a_J be the intermediate with the smallest number of occurrences: $w_J=\min_j(w_j)$

Let M_J be the set of new mechanisms which are constructed as follows. For each pair of mechanisms $m_k\in W_J$ and $m_b\in W_J$, where $m_k\neq m_b$:

Form a combination mechanism, m_c:

$m_c=\beta_{bJ}\, m_k-\beta_{kJ}\, m_b$

and its reverse:

$m_{-c}=\beta_{kJ}\, m_b-\beta_{bJ}\, m_k$

and compute the σ coefficients of m_c:

$\forall i=1, ...,S:\ \sigma_{ci}=\beta_{bJ}\,\sigma_{ki}-\beta_{kJ}\,\sigma_{bi}$

Consider all non-zero σ_{ci} coefficients for the directionality of the corresponding s_i steps. If there exists a step s_f such that $\rightarrow s_f$ and $\sigma_{cf}<0$, then m_c is rejected. If there exists a step s_f such that $\rightarrow s_f$ and $\sigma_{cf}>0$, then m_{-c} is rejected. If none of the two mechanisms is rejected

by these tests (i.e., if $\sigma_{cf}=0$ whenever $\rightarrow s_f$), then either one of the two mechanisms should be retained (but not both). If, for example, m_c is retained, we compute its β coefficients:

$\forall j \ni a_j \in (N \cup N_T)$: $\beta_{cj}=\beta_{bJ}\,\beta_{kj} - \beta_{kJ}\,\beta_{bj}$
and include m_c in the set of combination mechanisms M_J.

Phase 4. Update of active sets: The new set of active mechanisms is equal to $M:=(M \cup M_J)-W_J$ and the intermediate a_J is removed from N.

Phase 5. Check for termination: If $N \neq \emptyset$ then we go back to phase 2 (computation of numbers of combinations). If $N=\emptyset$ then the resulting set of mechanisms, M, is returned.

Parallelizability. The selection of intermediates need not be completely sequential. If the elimination of two intermediates involves disjoint sets of mechanisms, i.e., if $W_J \cap W_{J'}=\emptyset$, then a_J and $a_{J'}$ can be eliminated in parallel. This was done in Figure 1.3, 1.9, and 1.10. This can be generalized for any number of intermediates with disjoint W_j sets. A parallel implementation of the method is currently under development.

Incremental update of ranking of intermediates. While the number of combinations, n_j, is recomputed for each a_j after each iteration, many intermediates may not be affected at all by an iteration. Specifically, n_j will be unaffected by the elimination of a_J if $W_j \cap W_J=\emptyset$, i.e., precisely when a_j and a_J could have been processed (eliminated) in parallel. Thus, if a sequential implementation is used, one can selectively update n_j only if $W_j \cap W_J \neq \emptyset$.

Bookkeeping. Several choices in bookkeeping are only a matter of convenience and have no impact on the overall effectiveness of the algorithm.

A simplification is possible in keeping track of terminal species. It is not necessary to maintain the portion of the β_{kj} matrix that corresponds to terminal species, i.e., the columns a_{I+1} to a_A can be dropped from Figure 2.1 during the operation of the algorithm. They can be computed at the end, for all mechanisms in the final set M, from the matrices σ_{ki} and α_{ij}. This alternative, which was partly followed in Part 1 in Figure 1.9 for the construction of methanol mechanisms, can be stated as:

Variation 3 (Postponement of terminal species)

- In the initialization in **Phase 1** and in the formation of combinations in **Phase 4**, instead of computing the matrix entries for the new mechanisms for each species in the set $N \cup N_T$, the entries should be computed only for each intermediate in N.
- At the conclusion of Phase 5, computation of β_{kj} for j=I+1,...,A (i.e., for $a_J \in N_T$) is carried out:

$$\beta_{kj}=\sum_{i=1}^{S}\sigma_{ki}\alpha_{ij}.$$

Two other bookkeeping considerations will be mentioned here without formal or detailed description of the implied variations in the algorithm. First, the update of the σ entries can be postponed until the end, if all past mechanisms and their origin are stored even after they are rejected. At the very end, one can construct the σ entries of each mechanism from its origin. This is beneficial only if long mechanisms that are arise after some combinations only to be ultimately rejected. Note that this variation applies only to Algorithm 1, since Algorithms 2 and 3 need the σ entries for the elimination of redundant mechanisms. It is not clear whether this modification has practical value.

Second, whether mechanisms that have been rejected are deleted altogether depends on how much explanation one desires about how the algorithm reached its results. If no such explanation is needed then rejected mechanisms can be discarded and the origin column is unnecessary.

Implementational Decisions. In the conversion of the mathematical presentation of the algorithm into a computer program, many modifications can be made to enhance the program's efficiency. For example, as the matrices σ_{ki} and β_{kj} are likely to be very sparse for large systems, instead of simply implementing

them as arrays, one can use sparse-matrix implementation techniques which store only the non-zero elements of a matrix. This will reduce memory requirements and will make the time requirement of matrix operations dependent on the number of non-zero elements rather than the total number of elements.

In an object-oriented programming environment, one would implement species and mechanisms as objects, and would maintain (and updated as needed) pointers from each mechanism to all species that appear in its reaction stoichiometry, as well as from each species to all mechanisms whose reactions involve the species in question. Many operations in the algorithm are much more efficient when they use such pointers. Mavrovouniotis (1989) and Mavrovouniotis *et al.* (1990) employed this kind of implementation approach in their algorithm for the synthesis of biochemical pathways.

THE PROCEDURE OF HAPPEL AND SELLERS

This procedure, which has already been presented and applied in numerous publications (Happel and Sellers, 1982, 1983, 1989, Happel, 1986, Sellers, 1984, 1989, and Happel *et al.*, 1990), will be only briefly sketched here. In the description, an effort will be made to point out and remove superficial differences between their procedure and the alternative procedure described in this report, in order to highlight differences in the essential features of the two methods in the analysis of hypothetical chemical systems in a subsequent section.

1. Initialization. The initial setup of the Happel and Sellers procedure is shown in Figure 2.3 and is almost equivalent to that used by the method presented here in Figure 2.1 for the abstract, generalized setup of this report and in the examples in Part 1 (Figures 1.2 and 1.9). Essentially, Happel and Sellers regard each step as a partial mechanism, precisely the way our procedure does. The setup lists, for each mechanism m_k, the quantities σ_{ki} and β_{kj}, with the intermediate species having the indices j=1,..., I and the terminal species j=I+1,..., I+T (where I+T=A is the total number of species). In their own examples, Happel and Sellers do not list σ_{ki} in matrix form; they instead use an explicit linear combination of steps, of the form $m_k = \sum_{i=1}^{S} \sigma_{ki} s_i$, but this is clearly only a superficial difference.

The directionality labels used in our algorithm are not needed here, because each step is initially considered bidirectional. Mechanisms which violate the directionality restrictions on steps will be rejected at the end.

2. Diagonalization and identification of linearly independent mechanisms. The objective of this phase is the identification of linearly independent overall mechanisms and overall reactions. This is accomplished by converting the initial matrix of β_{kj} coefficients into the diagonal form of Figure 2.4 (without changing its size) and maintaining the linear independence of the σ rows of the matrix. To this end, one uses column and row operations which transform the partial mechanisms. Each row operation must be carried out identical for both the β and σ matrices. Happel and Sellers (1983) state that the σ rows of the final matrix will be linearly independent if one confines the operations to:

- Adding to (or subtracting from) a row a scalar multiple of a row lying *above* it; a multiple of any row k can thus be used to modify rows k+1 to S.
- Interchanging two columns (including their species labels) corresponding two intermediates, or two columns corresponding to terminal species; it is not permissible to interchange a terminal and an intermediate.
- Moving a row which contains only zeroes in its β portion (i.e., entails a null reaction) to the bottom of the matrix.

Although this description leaves some algorithmic details unspecified, there is no significant difficulty in implementing a specific procedure based on these rules. The procedure will resemble the well-known Gauss elimination method for solving linear systems. The computational burden that this procedure entails is rather small.

The diagonalization procedure converts the matrix of β_{kj} coefficients into the form of Figure 2.3 while maintaining the linear independence of the σ rows of the matrix. As shown in the figure, the procedure identifies a set of R linearly independent direct reactions, and one direct mechanism for each reaction. The

				a_1	a_2	.....	a_I	a_{I+1}	.....	a_A
σ_{11}	σ_{12}	.....	σ_{1S}	β_{11}	β_{12}	.....	β_{1I}	$\beta_{1(I+1)}$	.....	β_{1A}
σ_{21}	σ_{22}	.....	σ_{2S}	β_{21}	β_{22}	.....	β_{2I}	$\beta_{2(I+1)}$	.....	β_{2A}
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
σ_{S1}	σ_{S2}	.....	σ_{SS}	β_{S1}	β_{S2}	.....	β_{SI}	$\beta_{S(I+1)}$	.....	β_{SA}

Figure 23 Setup for the application of the algorithm of Happel and Sellers (1982, 1983, 1989). The left portion of the table contains the coefficients σ_{ki}; initially, each mechanism m_k contains only the corresponding step s_k, hence $\sigma_{ki}=1$ if $k=i$ while $\sigma_{ki}=0$ if $k\neq i$. The right portion of the table shows the reactions accomplished by the mechanisms. Species as a_1 to a_I are intermediates while a_{I+1} to a_A ($A=I+T$) are terminal species. Initially, $\beta_{kj}=\alpha_{ki}$.

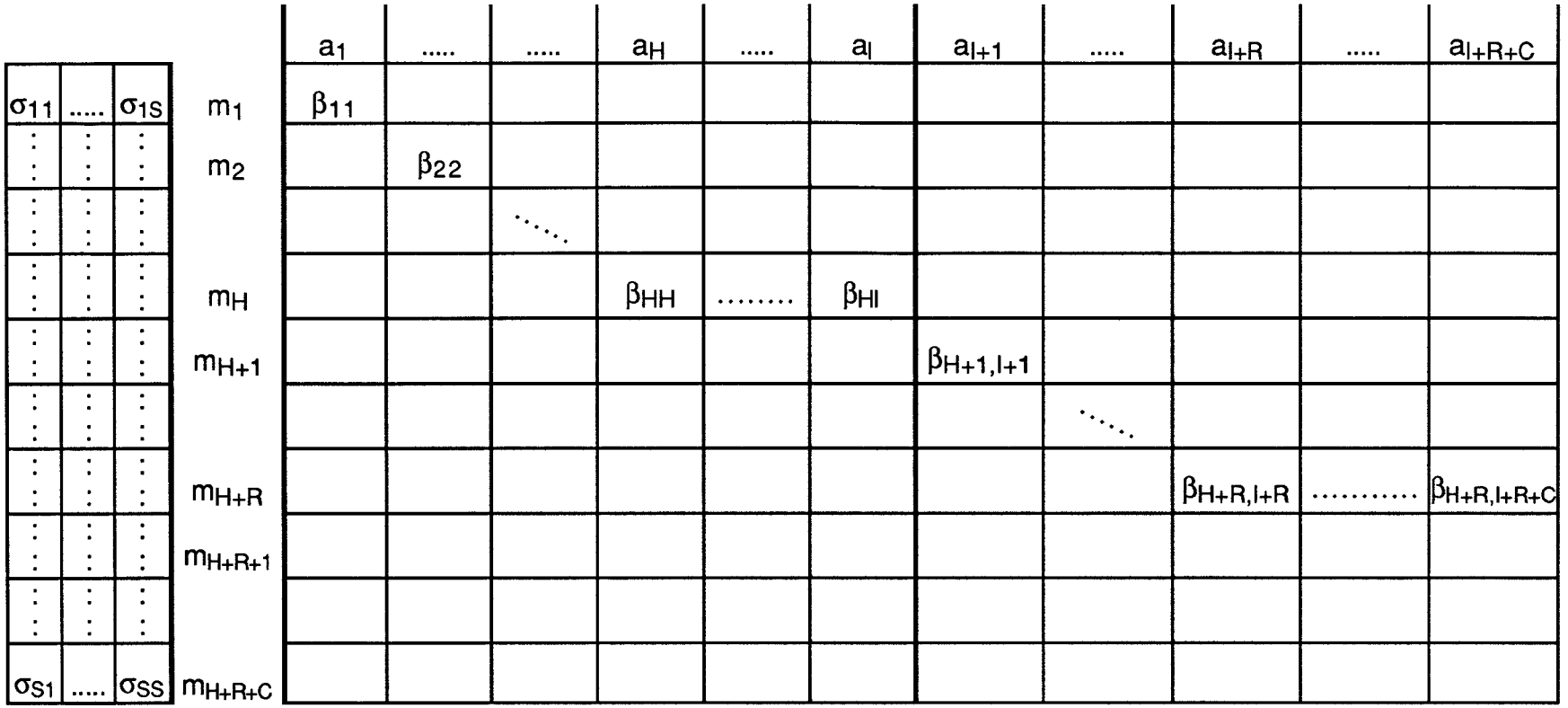


Figure 4. Diagonalized setup derived by the algorithm of Happel and Sellers (1982, 1983, 1989). The left portion of the table contains the coefficients σ_{ki}. The right portion of the table shows the reactions accomplished by the mechanisms. All entries below the marked diagonal are zero; entries on the marked diagonal are non-zero; and entries above the marked diagonal are not restricted. The reactions corresponding to the R mechanisms m_{H+1} to m_{H+R} form a basis for the set of overall reactions; we have one mechanism for each of these basis reactions. The C mechanisms m_{H+R+1} to m_{H+R+C} have null reactions and form a basis for the set of cyclic mechanisms. The R+C mechanisms m_H to m_{H+R+C} form a basis for the set of all mechanisms. Note that R+C=T (the number of terminal species), I+R+C=I+T=A (the total number of species), and H+R+C=S (the number of steps).

general overall reaction is a linear combination of the identified reactions. The procedure also constructs a set of C linearly independent cycles (whose overall reaction is null); these form a basis for the whole set of cycles of the chemical system. *Any* diagonalization that identifies R basis direct reactions (with one mechanism for each), and C linearly independent cycles is satisfactory.

3. Combination of linearly independent mechanisms. Direct mechanisms are formed from the R linearly independent non-null reaction mechanisms and the C linearly independent cycles, by considering each possible combination of C steps – and there are $\binom{S}{C}$ many such combinations – in conjunction with each of the non-null reaction mechanisms. The objective is to linearly combine the basis cycles and the non-cyclic mechanism into a new mechanism (for the same overall reaction) which does *not* include the chosen steps. Let **m** represent the vector of basis cycles and μ the vector of coefficients that will be used in the combination of the basis cycles; let m_1 be the chosen non-cyclic reaction mechanism. The new mechanism that will be constructed will be equal to $m_1+\mu\cdot\mathbf{m}$ and is determined as follows:

The C columns corresponding to the chosen steps, and the C rows of the linearly independent cycles in the left portion of Figure 2.4 define a C×C matrix denoted as **M** by Happel and Sellers (1982, 1983, 1989). The same C columns and the row corresponding to the chosen non-cyclic mechanism m_1 define a vector denoted by σ. The vector of mechanism coefficients, μ, is then obtained as the solution of the linear system:

$$\mathbf{M}\mu+\sigma=\mathbf{0} \Rightarrow \mu=-\mathbf{M}^{-1}\sigma$$

which defines precisely the requirement that the chosen C steps do not participate in the mechanism being constructed.

Four important elements remain for a complete description of the procedure. First, the matrix **M** might be singular; this simply means that it is *not* possible to eliminate the chosen C steps, and the combination is rejected.

Second, the above procedure allows duplicate mechanisms to be produced. Happel and Sellers (1983, p. 290) point out that this can be avoided by maintaining a listing of all the mechanisms as they are produced by the procedure, with the steps they involve. When a new combination of C out of S steps is chosen, the listing of the mechanisms is examined, to determine whether there already exists a mechanism which excludes these C steps; if so, then the combination is rejected without consideration of **M**.

Another issue is that of directionality of mechanism steps. The algorithm of Happel and Sellers does not take directionality into account in the construction of mechanisms. Thus, mechanisms which are infeasible because they violate the (thermodynamically or otherwise) postulated direction of a step are discarded after they are constructed. There is no apparent way to avoid the construction of such directionally infeasible mechanisms.

Finally, in the case of multiple overall reaction (i.e., when there are more than one choices for the non-cyclic m_0) a given combination of C steps gives one direct submechanism each choice of m_0, i.e., for each of the basis reactions. Combination of the submechanisms obtained for each basis reaction, under the same combination of C steps, provides the direct mechanism for the general overall reaction. Since the mechanism for the general overall reaction is actually obtained by combining submechanisms of individual basis reactions, it has the same number of degrees of freedom as the general overall reaction. This is illustrated in the application of the Happel and Sellers algorithm for a multiple-reaction hypothetical chemical system, in a subseqent section.

The construction of the mechanism of the general overall reaction must in fact precede the rejection of directionally infeasible mechanisms; this happens because the mechanism of the general reaction may permit elimination of the steps whose directionality appears to be violated in a particular component mechanism. For example, one of the basis reactions might entail directionally infeasible submechanisms, while its replacement with another (equally acceptable) reaction may remove the apparent infeasibility (Happel *et al.*, 1990, p.1061).

A HYPOTHETICAL SYSTEM WITH A SINGLE OVERALL REACTION

The properties of the the algorithms introduced here, as well as further improvements to these algorithms, will be discussed in a forthcoming report. It should be briefly mentioned here that the algorithm is both sound in generating only feasible overall mechanisms, and complete in ensuring that all direct mechanisms will be found. In the form of Algorithm 2 and especially Algorithm 3, the procedure is also non-redundant because it avoids duplicate and indirect mechanisms. The analysis of the behavior of the algorithm is aided significantly by the elegant mathematical theory of direct mechanisms presented by Sellers (1982, 1984) and Happel and Sellers (1982, 1983, 1989). It should be emphasized, however, that from the procedural viewpoint, i.e., in terms of the actual operation of the algorithms, there are very few similarities between the procedure recommended by Happel and Sellers (1982, 1983, 1989) and the procedure presented here.

The remainder of this report is devoted to a discussion of the algorithm in the context of model studies with abstract (hypothetical) chemical systems. These studies will permit particularly close examination of the effect of irreversible steps and the differences in the definition of direct mechanisms. The goal of these model studies is not to show superiority of one algorithm over the other but to raise some important issues and to dismiss any notion that the two algorithms are actually similar or only superficially different.

Chemical System Description. Consider the hypothetical chemical system depicted graphically in Figure 2.5, where A_1 through A_I are intermediates, while A_{I+1} and A_{I+2} are terminal species. To facilitate the analysis we take I to be an odd number; the parameter will be otherwise left unspecified and represents the size of the problem. The mechanism steps are:

step s_i, for i=1,..., I: $A_{I+1} \rightarrow A_i$ (steps s_1 to s_I)

step s_{I+i}, for i=1,..., I: $A_i \rightarrow A_{I+2}$ (steps s_{I+1} to s_{2I})

step s_{2I+i}, for i odd and $1 \le i \le I-2$: $A_i \rightarrow A_{i+1}$ (odd-numbered steps s_{2I+1} to s_{3I-2})

step s_{2I+i}, for i even and $2 \le i \le I-1$: $A_{i+1} \rightarrow A_i$ (even-numbered steps s_{2I+2} to s_{3I-1})

An important feature of this system is that even-numbered intermediates are produced through unidirectional steps from their neighboring odd-numbered intermediates. There are I intermediates, T=2 terminal species (hence A=I+2 total species), S=3I–1 steps, and only R=1 non-null overall reaction:

$A_{I+1} \rightarrow A_{I+2}$

Despite the apparent large number of steps this should be considered a sparse chemical system, because the number of steps is only linear in the number of species. If we confine our attention to steps of the form $A_i \rightarrow A_{i'}$, the total number of possible steps for I+2 species is equal to $(I+2)(I+1)/2=(I^2+3I+2)/2$, which is quadratic in I; the number of steps in the example considered here is only 3I–1, i.e., linear in I.

While this is a hypothetical system, it is useful to show here that there is nothing inherently absurd or chemically counterintuitive about a mechanism of this form. Consider as the initial reactant A_{I+1} the hydrocarbon depicted in Figure 2.6, with its secondary and tertiary carbons numbered as shown in the figure. In the initiation of fluid catalytic cracking reactions, a hydride is removed from the hydrocarbon, giving rise to an alkyl cation. Depending on which hydrogen is removed, the ion can be primary, secondary, or tertiary. In general, primary alkyl cations are unlikely to be formed. If we assume that both secondary and tertiary ions are possible, then we can denote as A_i (i=1 to I) the alkyl cation which bears the positive charge on the carbon numbered with i in Figure 2.6. Thus, the steps of the form $A_{I+1} \rightarrow A_i$ represent the formation of alkyl cations (neglecting the produced hydride); this is not a thermodynamically irreversible step, but it is considered unidirectional because it is expected to occur with a net forward rate. The interconversions of intermediates can be viewed as rearrangements of the alkyl cation through movement of a hydrogen; since tertiary ions are favored thermodynamically, the rearrangement is permitted only in the direction secondary→tertiary (i.e., odd-numbered intermediates are converted to even-numbered ones), which gives precisely the set of steps we have used in our hypothetical system for interconversion of intermediates. Finally, the conversion of any intermediate to the final product A_{I+2} corresponds to the virtually irreversible actual cracking (β-scission) of the molecule into smaller ones, with

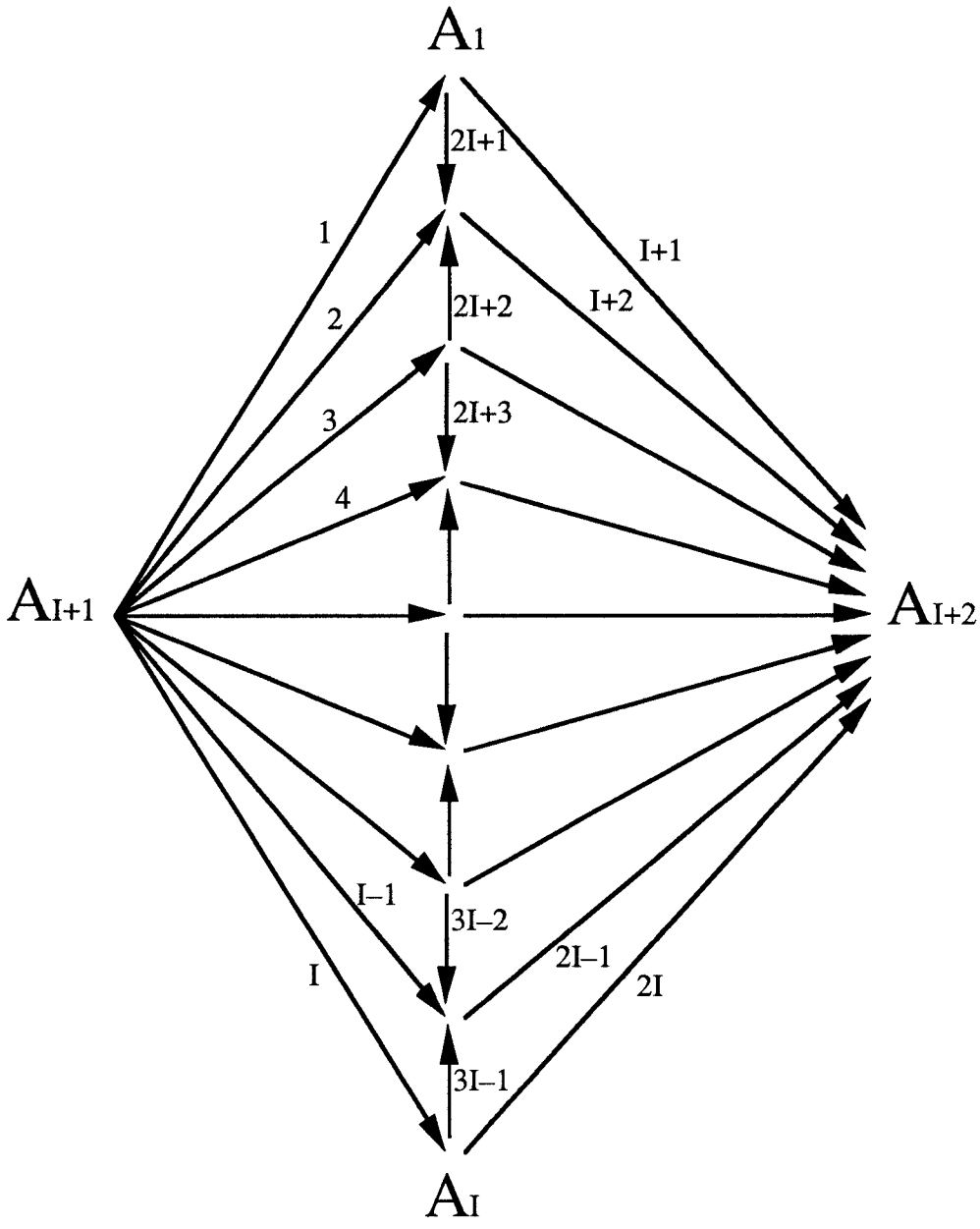


Figure 2.5. A hypothetical chemical system with unidirectional steps. The steps are numbered in the figure as 1, 2, ..., 3I–1. A_{I+1} and A_{I+2} are terminal species. The intermediates are A_1, A_2, ..., A_{I-1}, A_I (from top to bottom).

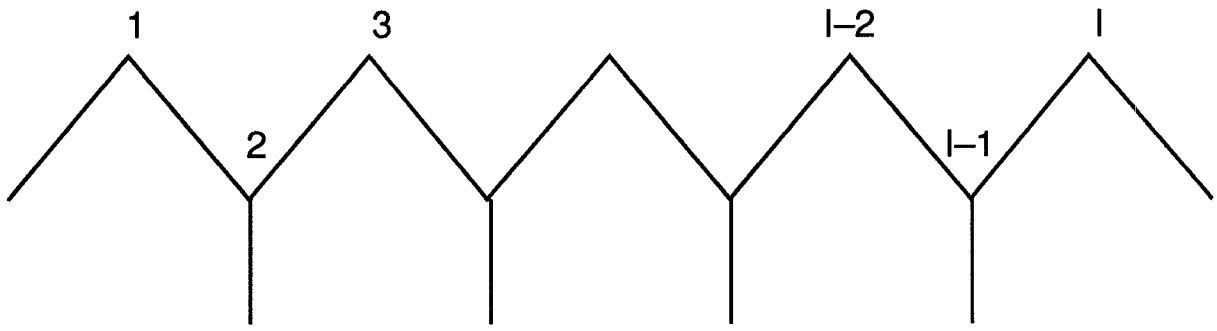


Figure 2.6. A hydrocarbon structure, with only its secondary and tertiary carbons numbered.

all possible products lumped into A_{I+2}; this kind of lumping makes sense if we are only interested in the conversion of A_{I+1} and not in product distribution.

This does not show, of course, that this particular chemical system is especially important; it does show that, although an idealized arrangement devised to analyze features of the algorithm, this hypothetical chemical system is *not* particularly unreasonable. The only exotic aspect of this chemical system is its excessive symmetry or regularity. The regularity is necessary for counting mechanisms and cycles and analyzing the behavior of the algorithms. If some of the symmetry of the system is lost, precise counting becomes more difficult but the general trends, and consequently the qualitative conclusions, still hold.

Direct Mechanisms through the Proposed Algorithm. Let us consider now the number of directionally-feasible direct mechanisms in this system. Each even-numbered intermediate cannot be converted to another intermediate. Each odd-numbered intermediate can be converted either directly to the final product or first into its neighboring even-numbered intermediates and then into the final product. There are, thus, three kinds of direct mechanisms (Figure 2.7):

(a) m_i, for i=1,...,I: $A_{I+1} \rightarrow A_i \rightarrow A_{I+2}$

(b) m_{I+i}, for i odd and $1 \le i \le I-2$: $A_{I+1} \rightarrow A_i \rightarrow A_{i+1} \rightarrow A_{I+2}$

(c) m_{I+i}, for i even and $2 \le i \le I-1$: $A_{I+1} \rightarrow A_{i+1} \rightarrow A_i \rightarrow A_{I+2}$

There are I mechanisms of type (a), and (I–1)/2 mechanisms of each of the types (b) and (c). Hence, there are 2I–1 direct mechanisms in total.

How would the procedure described in this report construct these mechanisms? We first consider each step to be a partial mechanism, and process the intermediates as prescribed by the algorithm. The first two intermediates to be processed are A_1 and A_I (in either order). They would lead to the direct mechanisms $A_{I+1} \rightarrow A_1 \rightarrow A_{I+2}$ and $A_{I+1} \rightarrow A_I \rightarrow A_{I+2}$ as well as two additional partial mechanisms, $A_{I+1} \rightarrow A_1 \rightarrow A_2$ and $A_{I+1} \rightarrow A_I \rightarrow A_{I-1}$; they would also abolish the steps $A_{I+1} \rightarrow A_1$, $A_1 \rightarrow A_{I+2}$, $A_1 \rightarrow A_2$, $A_{I+1} \rightarrow A_I$, $A_I \rightarrow A_{I-1}$, $A_I \rightarrow A_{I+2}$.

From this point on, all the intermediates will have $n_j=3$ and $w_j=4$; they will be tied even if Variations 1 or 2, or Algorithm 4 are used. Thus, the order of processing is arbitrary; one could have even processed all of the intermediates (including A_1 and A_I) in an arbitrary order and would still obtain the same results and expend the same computational effort. In this system, *no redundant mechanisms are ever constructed*, i.e., Algorithms 2 and 3 play no role here; the basic Algorithm 1 would be sufficient. Apart from the one-step initial mechanisms, *no mechanism is ever discarded.* Thus, no computational effort is wasted.

Each intermediate from A_2 to A_{I-1} requires the construction of 3 mechanisms. Thus, the total number of mechanism-construction operations that will be performed is equal to 3I–2 (the –2 originating from the fact that A_1 and A_I require only two mechanism-constructions each). This can be related to the number of direct mechanisms, 2I–1, as follows. Each of the I mechanisms of type (a) requires only one mechanism-combination operation, since it consists of only two steps. Each of the I–1 mechanisms of type (b) and (c) consists of three steps; since we combine two partial mechanisms at a time, we need two mechanism-combination operations for each of these mechanisms. The total is I+2(I–1)=3I–2 operations.

Mechanisms through the Happel and Sellers Algorithm. Consider, now, the operation of the Happel and Sellers (1982, 1983, 1989) algorithm on this example. Each step is initially treated as reversible. There is only one overall reaction, and any one mechanism for it, e.g., $A_{I+1} \rightarrow A_1 \rightarrow A_{I+2}$, will be included in the basis. The basis will also include, however, a total of C=2I–2 linearly independent cyclic mechanisms. There is one basis which includes only the shortest possible cyclic mechanisms. These are:

(a) m_i, $1 \le i \le I-1$: $A_{I+1} \rightarrow A_i \rightarrow A_{i+1} \rightarrow A_{I+1}$

(b) m_{I-1+i}, $1 \le i \le I-1$: $A_{I+2} \rightarrow A_i \rightarrow A_{i+1} \rightarrow A_{I+2}$

There is, of course no guarantee that this particular set of cyclic mechanisms will be used, but the analysis is independent of the choice of basis. Note that each cyclic mechanism uses at least one step in the wrong

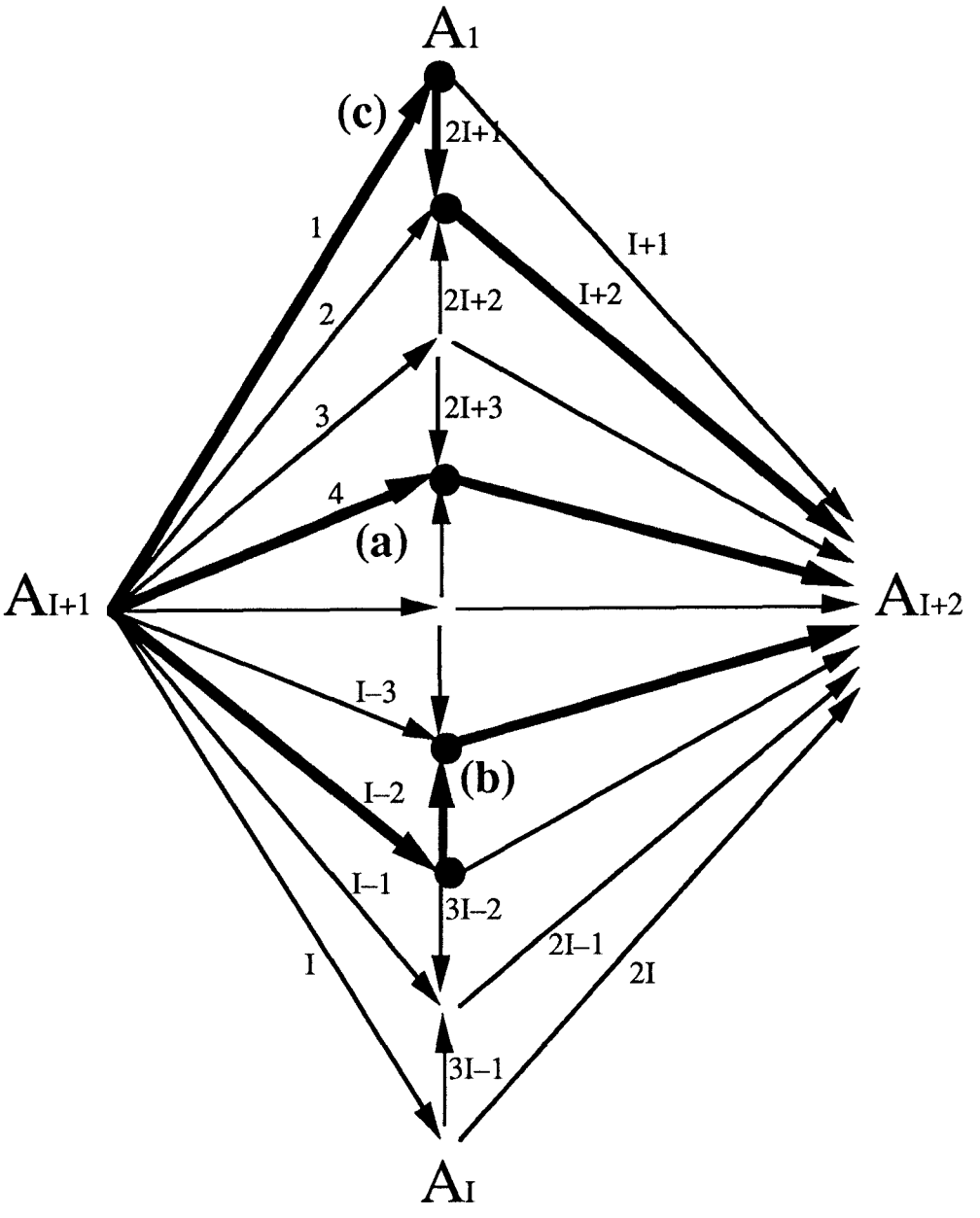


Figure 2.7. The three classes of direct mechanisms for the hypothetical chemical system of Figure 2.5. Mechanism (a) can be written as $A_{I+1} \rightarrow A_4 \rightarrow A_{I+2}$ and is of general type $A_{I+1} \rightarrow A_i \rightarrow A_{I+2}$ (i=1,...,I); mechanism (b) can be written as $A_{I+1} \rightarrow A_{I-2} \rightarrow A_{I-3} \rightarrow A_{I+2}$ and is of the general type $A_{I+1} \rightarrow A_i \rightarrow A_{i+1} \rightarrow A_{I+2}$ (for i odd and $1 \leq i \leq I-2$); mechanism (c) can be written as $A_{I+1} \rightarrow A_1 \rightarrow A_2 \rightarrow A_{I+2}$ and is of the general type $A_{I+1} \rightarrow A_{i+1} \rightarrow A_i \rightarrow A_{I+2}$ (for i even and $2 \leq i \leq I-1$).

direction; for example, a cyclic mechanism of type (a) always uses the step $A_{i+1} \rightarrow A_{I+1}$ in the wrong direction (and if i is even, then the mechanism also uses $A_i \rightarrow A_{i+1}$ in the wrong direction). In the procedure of Happel and Sellers, however, this is not prohibited; only the final mechanisms can be rejected on directionality grounds. This issue will be reexamined at the end of this example.

As Sellers (1989, p. 312) points out, there is a factor of $\binom{S}{C}$ in the computational effort required by the algorithm, because each combination of C (out of S) steps will be considered. Here, we have a number of combinations equal to:

$$\binom{S}{C} = \binom{3I-1}{2I-2} = \binom{3I-1}{I+1} = \frac{(3I-1)(3I-2)...(2I-1)}{(I+1)I(I-1)...2}$$

This quantity grows extremely fast. One can show, for example, that:

$$\binom{S}{C} > 3^{I+1} \quad \text{for } I \geq 5$$

$$\binom{S}{C} > I^{I/2} \quad \text{for } 3 \leq I \leq 39$$

Some sample values for $\binom{S}{C}$ are:

$$\binom{S}{C} \approx 5.3 \times 10^{6} \quad \text{for I=9}$$

$$\binom{S}{C} \approx 9.7 \times 10^{9} \quad \text{for I=13}$$

$$\binom{S}{C} \approx 3.4 \times 10^{16} \quad \text{for I=21}$$

For some of these combinations, one will need to invert a C×C matrix – a (2I–2)×(2I–2) matrix in this case – or determine that the matrix is singular, as discussed in the description of the Happel and Sellers algorithm given in an earlier section; this is another computationally costly operation. For other combinations this will not be necessary; specifically, if a mechanism has already been constructed which does not use the C (=2I–2) steps of the combination being considered. However, each of the $\binom{S}{C}$ combinations must be considered (Happel and Sellers, 1983, p. 290-291, Sellers, 1989, p. 312).

Direct Mechanisms produced by Happel and Sellers, Disregarding Directionality. The number of combinations that will require, and lead to, successful matrix inversions is equal to the number of direct mechanisms in this chemical system. However, since this example involves unidirectional steps, the Happel and Sellers algorithm will construct not only the directionally-feasible direct mechanisms described earlier, but also other mechanisms which would be feasible only if all steps were bidirectional; essentially, the algorithm will construct mechanisms for the system of Figure 2.8.

How many direct mechanisms are there, in this example, if we disregard the directionality of steps? There are three types of direct mechanisms, shown in Figure 2.9:

(a) For all i ($1 \leq i \leq I$): $A_{I+1} \rightarrow A_i \rightarrow A_{I+1}$

(b) For all i and all u ($1 \leq i < u \leq I$): $A_{I+1} \rightarrow A_i \rightarrow A_{i+1} \rightarrow ... \rightarrow A_{u-1} \rightarrow A_u \rightarrow A_{I+1}$

(c) For all i and all u ($1 \leq u < i \leq I$): $A_{I+1} \rightarrow A_i \rightarrow A_{i-1} \rightarrow ... \rightarrow A_{u+1} \rightarrow A_u \rightarrow A_{I+1}$

These types correspond to the three types of direct mechanisms listed earlier for the unidirectional system, except that in types (b) and (c) neglecting the directionality steps allows a sequence of more than two intermediates in the mechanism. There are I mechanisms for type (a). For mechanisms of type (b), there are I ways to choose i, and I–i ways to choose u, giving a total number of mechanisms of type (b):

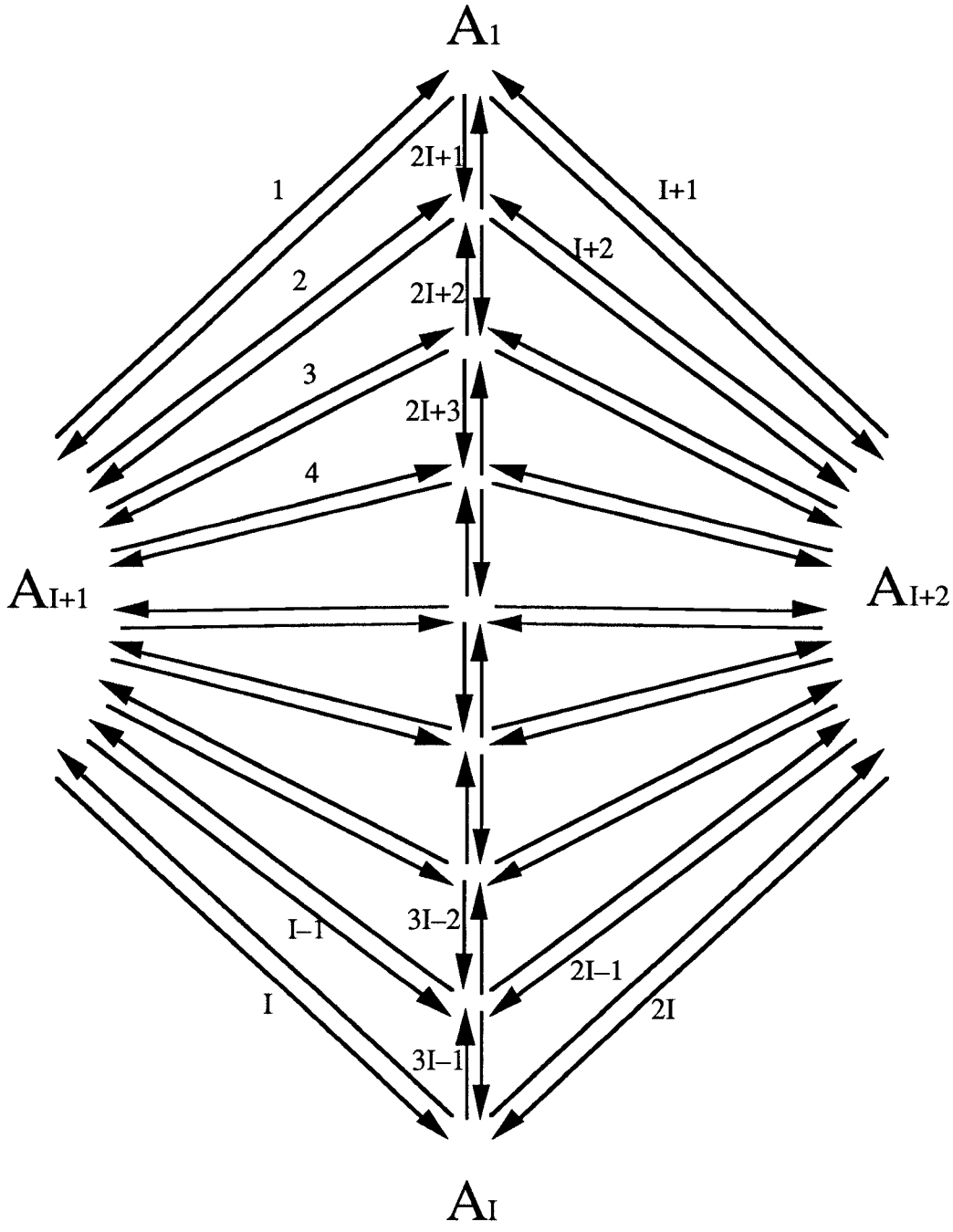


Figure 2.8. A modification of the directionality of the chemical system of Figure 2.5. Here, all steps are taken to be bidirectional.

$$\sum_{i=1}^{I}(I-i) = \sum_{i=1}^{I} I - \sum_{i=1}^{I} i = I^2-I(I+1)/2=I(I-1)/2$$

For mechanisms of type (c), there are I ways to choose i and i–1 ways to choose u, hence

$$\sum_{i=1}^{I}(i-1) = \sum_{i=1}^{I} i - \sum_{i=1}^{I} 1 = I(I+1)/2-I = I(I-1)/2$$

The total number of mechanisms is thus:

$I+I(I-1)/2+I(I-1)/2=I^2$

The three types of mechanisms can also be collectively considered to be of the form:

$$A_{I+1} \leftrightarrows A_i \leftrightarrows A_{i+1} \leftrightarrows \ldots \leftrightarrows A_{u-1} \leftrightarrows A_u \leftrightarrows A_{I+1}$$

where i and u can be either identical or different. The existence of I^2 direct mechanisms is a consequence of the fact that there are I^2 ways to pick an ordered pair of (not necessarily distinct) intermediates, (A_i, A_u); if the two intermediates are identical (i=u), we obtain a mechanism of type (a); if the first intermediate has a lower index than the second (i.e., i<u), a mechanism of type (b); and if i>u, a mechanism of type (c).

In this example, then, there are I^2 direct mechanisms if one ignores the directionality of steps, but only 2I–1 directionally feasible mechanisms. The Happel and Sellers algorithm must construct all I^2 mechanisms and only then eliminate those that violate directionality restrictions. For each of these I^2 mechanisms, a (2I–2)×(2I–2) matrix must be inverted.

Singular Matrices. A large number of combinations will lead to a singular matrix; specifically, if the choice of 2I–2 steps includes at least one step from each direct mechanism, the matrix **M** will be singular because no mechanism can be derived from the remaining steps. It is difficult to count the total number of such combinations, and we confine our attention to a subset of these.

If all steps $s_1, s_2, \ldots, s_I$ are included in the combination, then no direct mechanism is possible, because no step consuming A_{I+1} remains. Thus, any choice of I–2 out of the remaining 2I–1 steps will lead to a singular matrix. A similar argument applies to the symmetric case in which steps $s_{I+1}, s_{I+2}, \ldots, s_{2I}$ are included in the combination and A_{I+2} cannot be produced. Thus, the number of combinations that lead to a singular matrix is greater than:

$$2\binom{2I-1}{I-2} = 2\frac{(2I-1)(2I-2)\ldots(I+2)}{(I-2)(I-3)\ldots 2}$$

It can be shown that:

$$2\binom{2I-1}{I-2} > 3^I \quad \text{for } I\geq 7$$

and some values showing how fast this number of combinations grows are:

$$2\binom{2I-1}{I-2} \approx 3.9\times10^4 \quad \text{for } I=9$$

$$2\binom{2I-1}{I-2} \approx 8.9\times10^6 \quad \text{for } I=13$$

$$2\binom{2I-1}{I-2} \approx 4.9\times10^{11} \quad \text{for } I=21$$

As was noted in the derivation, this is merely a *lower bound* on the number of combinations that lead to singular matrices **M**. The *actual* number is likely to be higher but exhibit a similar rate of growth.

Modification of Directionality. It was noted earlier that each cyclic mechanism in the basis used by the Happel and Sellers algorithm for this example uses at least one step in the wrong direction. One may wonder, then, whether pruning of directionally infeasible mechanisms in the basis could avert the

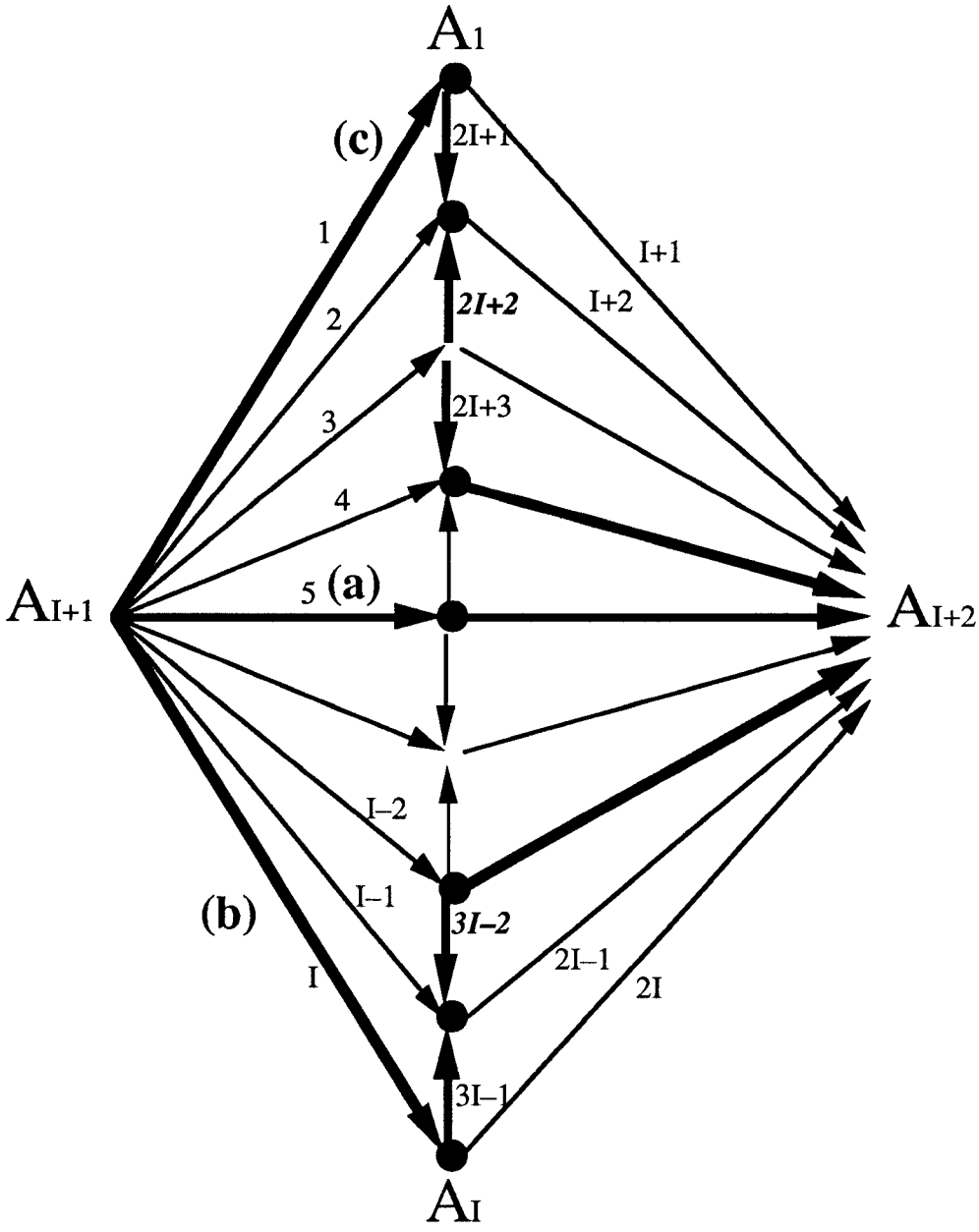


Figure 2.9. The three classes of direct mechanisms for the hypothetical chemical system of Figure 2.5, if the directionality of the steps is neglected (as in Figure 2.8). Mechanism (a) is of the general type $A_{I+1} \rightarrow A_i \rightarrow A_{I+1}$ $(1 \le i \le I)$, mechanism (b) of the type $A_{I+1} \rightarrow A_i \rightarrow A_{i+1} \rightarrow ... \rightarrow A_{u-1} \rightarrow A_u \rightarrow A_{I+1}$ $(1 \le i < u \le I)$, and mechanism (c) of the general type $A_{I+1} \rightarrow A_i \rightarrow A_{i-1} \rightarrow ... \rightarrow A_{u+1} \rightarrow A_u \rightarrow A_{I+1}$ $(1 \le u < i \le I)$. Note that the italicized steps 2I+2 and 3I–2 are actually used by the mechanisms in the direction opposite to that originally assumed. The three types of mechanisms correspond to those of Figure 2.7, except that for types (b) and (c) a sequence of more than two intermediates may occur in the mechanism.

combinatorial explosion. However, Happel *et al.* (1990) state clearly that thermodynamic feasibility or other directionality considerations must come at the very end, after the construction of the mechanisms, and this is the course they follow in their examples.

In this example, if one insists that the cyclic mechanisms in the basis be directionally feasible, *no feasible cycle* can be found, and no basis could be constructed. This is a consequence of the fact that a cycle must include either A_{I+1} or A_{I+2} (the two terminal species); all steps in which A_{I+1} participates have directions that cause consumption and never production of A_{I+1}, and similarly all steps in which A_{I+2} participates have directions that cause production and never consumption of A_{I+2}. This argument applies not only to the cycles used in the application of the Happel and Sellers algorithm above, but also to all other cycles in this chemical system.

The fact that directionality of cycles in the basis cannot be used to decide directionality of the final mechanisms will be shown here in another way, by modifying the directionality restrictions in this example. Specifically, the even-numbered steps s_2 to s_{I-1}, and the odd-numbered steps s_{I+2} to s_{2I-1} will be made bidirectional, leading to the following system (Figure 2.10):

step s_i, for i even and $2 \le i \le I-1$: $A_{I+1} \leftrightarrows A_i$

step s_i, for i odd and $1 \le i \le I$: $A_{I+1} \rightarrow A_i$

step s_{I+i}, for i even and $2 \le i \le I-1$: $A_i \leftrightarrows A_{I+2}$

step s_{I+i}, for i odd and $1 \le i \le I$: $A_i \rightarrow A_{I+2}$

step s_{2I+i}, for i even and $2 \le i \le I-1$: $A_{i+1} \rightarrow A_i$

step s_{2I+i}, for i odd and $1 \le i \le I-2$: $A_i \rightarrow A_{i+1}$

This system has a basis set of cyclic mechanisms that are directionally feasible (Figure 2.11).

(a) for i odd and $1 \le i \le I-2$: $A_{I+1} \rightarrow A_i \rightarrow A_{i+1} \leftrightarrows A_{I+1}$

(b) for i even and $2 \le i \le I-1$: $A_{I+1} \rightarrow A_{i+1} \rightarrow A_i \leftrightarrows A_{I+1}$

(c) for i odd and $1 \le i \le I-2$: $A_{I+1} \rightarrow A_i \rightarrow A_{I+2} \leftrightarrows A_{i+1} \leftrightarrows A_{I+1}$

(d) for i even and $2 \le i \le I-1$: $A_{I+1} \rightarrow A_{i+1} \rightarrow A_{I+2} \leftrightarrows A_i \leftrightarrows A_{I+1}$

This modification leaves the set of directionally feasible direct mechanisms unchanged. The analysis carried out above, including the determination of number of mechanisms that the Happel and Sellers algorithm must construct, as well as the number of cases leading to a singular matrix, remains valid. Thus, a basis of directionally feasible cycles still leads to the construction of directionally infeasible direct mechanisms that must be rejected in the end.

The Procedure of Milner. A definition of direct mechanisms was first given by Milner (1964), who pointed out that, for a system with I intermediates and a single overall reaction, a direct mechanism can include no more than I+1 steps with non-zero coefficients σ_{ki}. Thus, Milner (1964) proposed that direct mechanisms can be constructed by examining each combination of I+1 steps, and determining the σ_{ki} so that the intermediates do not appear in the overall stoichiometry and the overall reaction is accomplished.

In the example considered here, this means examining $\binom{3I-1}{I+1}$ combinations of steps. A basic property of binomial coefficients is that:

$$\binom{S}{C} = \binom{S}{S-C}$$

Hence:

$$\binom{3I-1}{I+1} = \binom{3I-1}{2I-2}$$

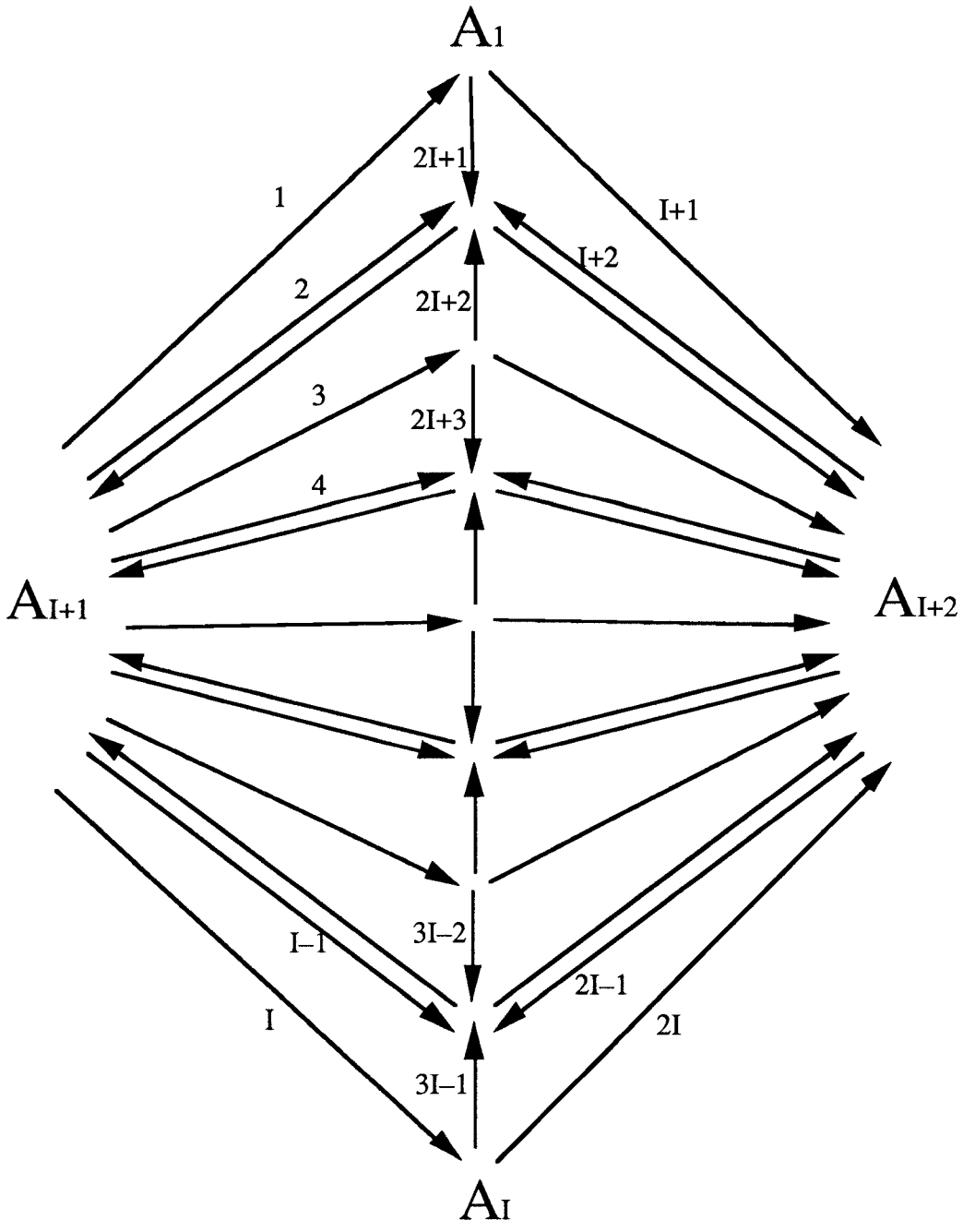


Figure 2.10. A modification of the directionality of the chemical system of Figure 2.5. Here, the even-numbered steps that involve either of the terminal species, A_{I+1} or A_{I+2}, are bidirectional, given that I is assumed odd; these are steps s_i: $A_{I+1} \leftrightarrows A_i$ for i even and $2 \le i \le I-1$, and s_{I+i}: $A_i \leftrightarrows A_{I+2}$ for i odd and $I+2 \le i \le 2I-1$. The remaining steps, and all steps that involve only intermediates, are unidirectional as was shown in Figure 2.5.

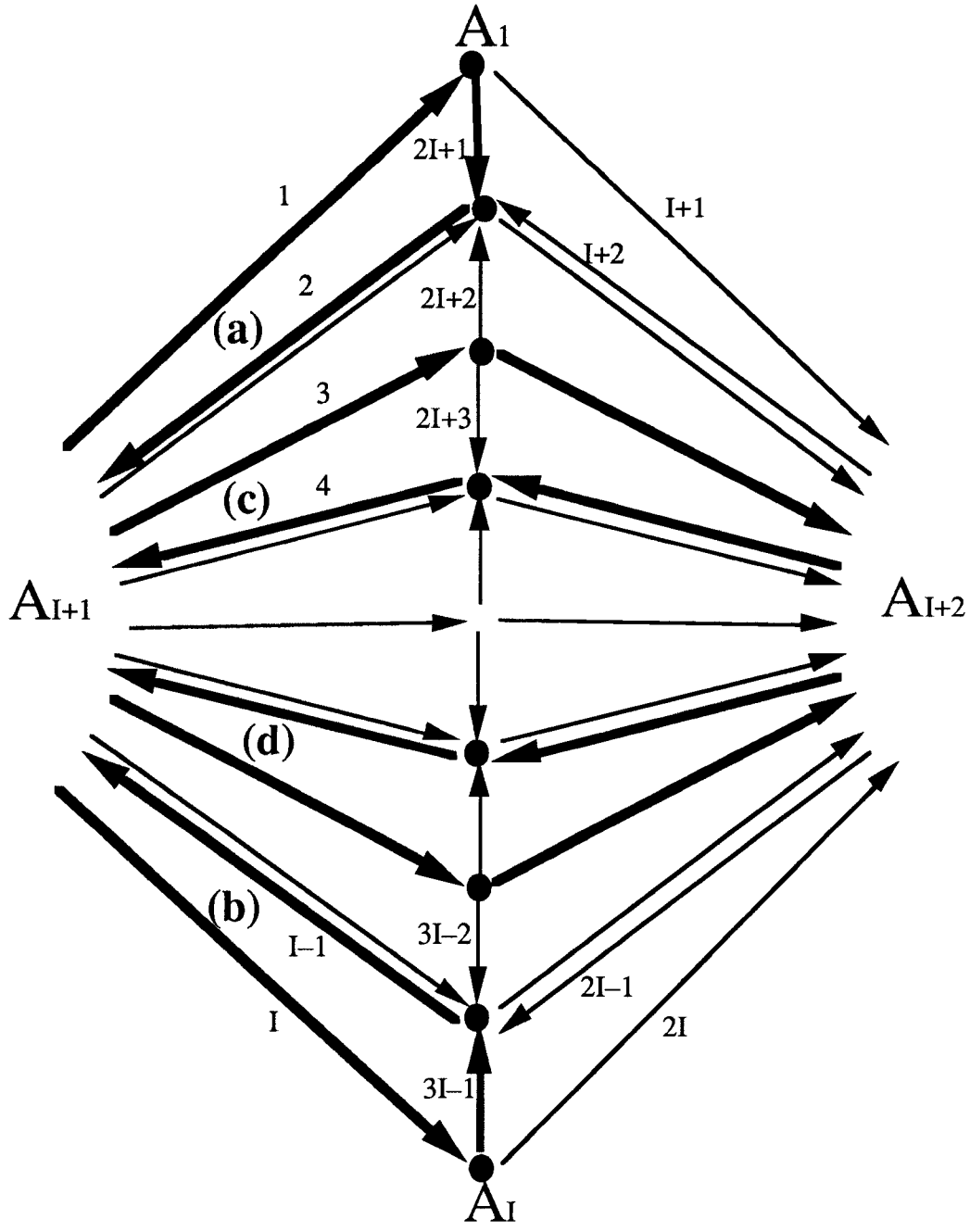


Figure 2.11. For the chemical system of Figure 2.10, there exists a basis set of cyclic mechanisms that are directionally feasible. The four types of cyclic mechanisms used in the basis are depicted here. Mechanism (a), $A_{I+1}\rightarrow A_1\rightarrow A_2\leftrightarrows A_{I+1}$ is of the general type $A_{I+1}\rightarrow A_i\rightarrow A_{i+1}\leftrightarrows A_{I+1}$ (for i odd and $1\leq i\leq I-2$). Mechanism (b), $A_{I+1}\rightarrow A_I\rightarrow A_{I-1}\leftrightarrows A_{I+1}$ is of the general type $A_{I+1}\rightarrow A_{i+1}\rightarrow A_i\leftrightarrows A_{I+1}$ (for i even and $2\leq i\leq I-1$). Mechanism (c), $A_{I+1}\rightarrow A_3\rightarrow A_{I+2}\leftrightarrows A_4\leftrightarrows A_{I+1}$ is of the general type $A_{I+1}\rightarrow A_i\rightarrow A_{I+2}\leftrightarrows A_{i+1}\leftrightarrows A_{I+1}$ (for i odd and $1\leq i\leq I-2$). Mechanism (d), $A_{I+1}\rightarrow A_{I-2}\rightarrow A_{I+2}\leftrightarrows A_{I-3}\leftrightarrows A_{I+1}$ is of the general type $A_{I+1}\rightarrow A_{i+1}\rightarrow A_{I+2}\leftrightarrows A_i\leftrightarrows A_{I+1}$ (for i even and $2\leq i\leq I-1$).

which means that the same number of combinations that must be considered by the procedure of Milner (1964) and the procedure of Happel and Sellers (1982, 1983, 1989). This is a consequence of the fairly large number of cycles present in this example. If the system contained the same number of intermediates and steps but fewer cycles, the procedure of Happel and Sellers is more efficient; this situation is generally likely in practice. It should also be noted that Milner (1964) did not explicitly addressed systems with multiple reactions.

A SYSTEM WITH A MULTIPLE OVERALL REACTION

The issue of the definition of direct mechanisms is examined here in the context of a very small system in which more than one overall reaction is possible. In systems with multiple overall reactions, it should be remembered that Happel and Sellers construct direct mechanisms for particular reactions. In fact, they define a direct mechanism in terms of the particular reaction it accomplishes: A mechanism m_k is direct if no other mechanism *for the same reaction* uses a subset of the steps used by m_k. Thus, when referring to sets of direct mechanisms we should be careful to name the reaction for which the mechanisms are direct. On the other hand, the algorithm proposed here does not look at the issue of direct reactions and linearly independent reactions. It considers a mechanism m_k as direct, regardless of the reaction which it accomplishes, if no other *overall* mechanism uses a subset of the steps used by m_k.

Chemical System Description. The chemical system shown in Figure 2.12, involves six steps, three intermediates, and three terminal species (S=6, I=3, T=3, A=6). The steps are

s_1: $A_4 \leftrightarrows A_1$

s_2: $A_1 \leftrightarrows A_5$

s_3: $A_5 \leftrightarrows A_2$

s_4: $A_2 \leftrightarrows A_6$

s_5: $A_6 \leftrightarrows A_3$

s_6: $A_3 \leftrightarrows A_4$

and they are all assumed bidirectional, since the effects of directionality have been discussed in a previous example.

Three direct overall reactions are possible in this system,

r_1: $A_4 \leftrightarrows A_5$

r_2: $A_5 \leftrightarrows A_6$

r_3: $A_6 \leftrightarrows A_4$

but only two (any two) are linearly independent. The general overall reaction can be written, for example, as a linear combination of r_1 and r_2:

$$r = \theta r_1 + (1-\theta) r_2$$

giving:

$$r: \; -\theta A_4 + (2\theta - 1) A_5 + (1-\theta) A_5 \leftrightarrows 0$$

Setting $\theta=0$, this expression gives r_2; setting $\theta=1$, it gives r_1; and setting $\theta=-1$, the expression gives r_3.

Happel and Sellers Algorithm. For systems with multiple overall reactions, the diagonalization procedure of Happel and Sellers (1982, 1983, 1989) will identify linearly independent overall reactions. Here, the values of the parameters mentioned in the diagonalization (Figure 2.4) are R=2 and C=1.

The reactions r_1 and r_2 will be arbitrarily taken as the basis reactions. The procedure will also identify one mechanism for each of the two basis reactions, and the one cyclic mechanism. We will assume that the shortest mechanisms of r_1 and r_2 were the ones constructed:

m'_1: s_1+s_2

m'_2: s_3+s_4

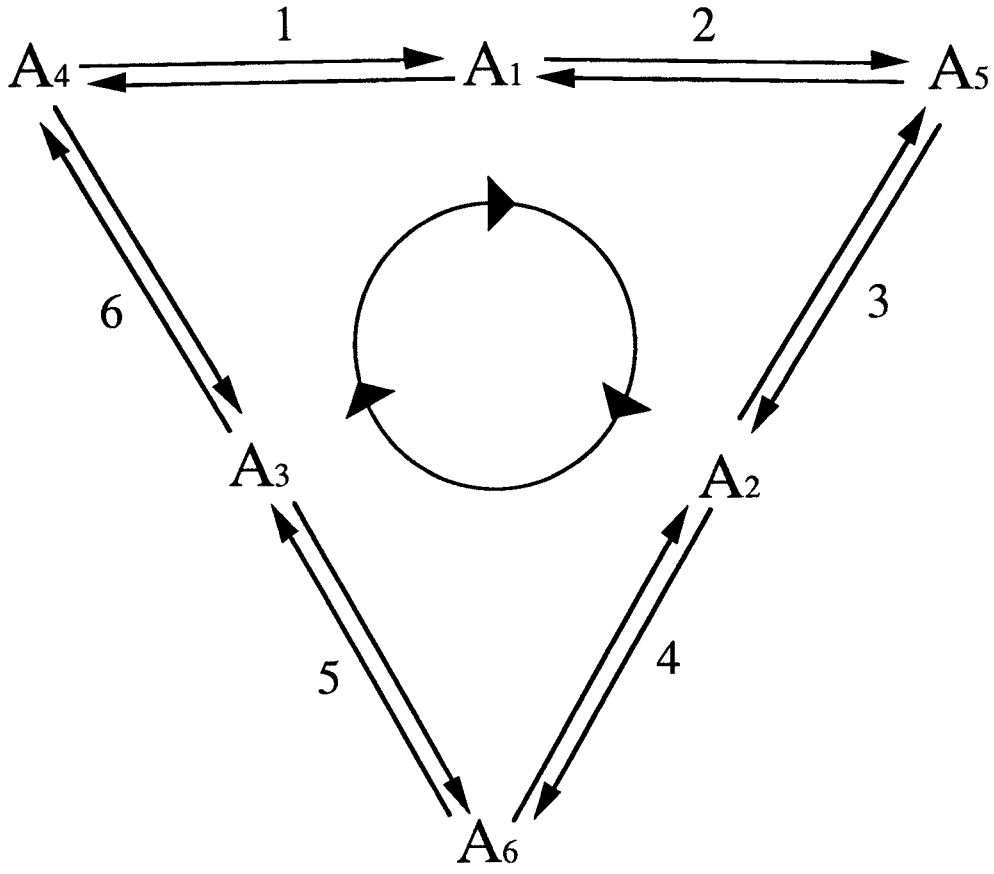


Figure 2.12. A small system with a multiple overall reaction. A_1, A_2, and A_3 are intermediates (I=3), while A_4, A_5, and A_6 are terminal species (T=3, A=I+T=6). There are six bidirectional steps, labelled in the figure. The arrows in the circle, the clockwise direction, indicate the forward direction for the mechanism steps.

The cyclic mechanism is unique:

m'_o: $s_1+s_2+s_3+s_4+s_5+s_6$

These basis reactions and mechanisms correspond to a diagonalization whose useful portion is shown in Figure 2.13. The general (indirect) mechanism m for the general overall reaction can be viewed as a combination of the linearly independent mechanisms m'_1, m'_2, and m'_o:

$$m=\theta m'_1+(1-\theta)m'_2+\phi m'_o$$

$$m=(\theta+\phi)\, s_1 + (\theta+\phi)\, s_2 + (1-\theta+\phi)\, s_3 + (1-\theta+\phi)\, s_4 + \phi\, s_5 + \phi\, s_6$$

The basis set of reactions, r_1 and r_2, although not unique (e.g., the equally acceptable set of r_2 and r_3 could be constructed by the algorithm) is quite useful because it provides an expression for the general overall reaction. The basis mechanisms m'_1, m'_2, and m'_o which were identified along with r_1 and r_2 during the diagonalization phase are also useful for constructing an expression for the general (indirect) overall mechanism, as was done above. However, they give no immediate information about, and bare no close relationship to, direct mechanisms. In fact, the diagonalization phase of the Happel and Sellers algorithm is concerned only with linear independence of reactions and mechanisms and can be fully understood and utilized without allusion to the concept of direct mechanisms.

The consideration of direct mechanisms is the main concern in the last phase of the algorithm, which considers each of the $\binom{S}{C}$ possible combinations of steps, and combines each of the basic non-cyclic mechanisms with all the cyclic mechanisms to eliminate the selected combination of steps. Here, $\binom{S}{C}=\binom{6}{1}=6$, i.e., we select one step (out of six) at a time.

Suppose s_1 is selected and let m_1 denote the resulting mechanism. To identify m_1, a linear combination of m'_1 and m'_o which eliminates s_1 will be constructed, and then a similar linear combination of m'_2 and m'_o. Happel and Sellers (1989) provide general formulae for this construction. In the small chemical system considered here, inspection of Figure 2.13 reveals that the correct linear combinations are shown here:

- Combining m'_1 and m'_o gives $m'_1-m'_o$ as the form of the direct mechanism m_1 for r_1. The mechanism can be written as a combination of steps as $-s_3-s_4-s_5-s_6$.
- Combining m'_2 and m'_o gives $m'_2+0m'_o=m'_2$ $(=s_3+s_4)$ as the form of the direct mechanism m_1 for r_2.

The combination of the two mechanisms constructed above produces the mechanism m_1 for the general reaction $r=\theta r_1+\phi r_2$:

m_1: $\theta(-s_3-s_4-s_5-s_6)+(1-\theta)(s_3+s_4)$

or

m_1: $(1-2\theta)s_3+(1-2\theta)s_4-\theta s_5-\theta s_6$

Selection of s_2 would yield the same mechanisms; as Happel and Sellers (1983, p. 290-291) show, the selection can be immediately rejected because the mechanisms constructed above already exclude s_2.

The submechanisms m_1 for r_1 (line 1 in Figure 2.14) and m_1 for r_2 (line 2 in Figure 2.14) are indeed direct mechanisms for their respective reactions. However, the arbitrary choice of r_1 and r_2 limit the significance of these direct submechanisms; note also that m_1 for r_1 (equal to $-s_3-s_4-s_5-s_6$) is particularly counterintuitive because it can be decomposed into the simpler submechanisms $-s_3-s_4$ (a direct mechanism for r_2) and $-s_5-s_6$ (a direct mechanism for r_3). The real significance of the submechanisms of m_1 for r_1 and r_2 is that their combination gives the mechanism m_1 for the general reaction r (line 3 of Figure 2.14) which is a direct mechanism for the general overall reaction.

Selection of s_3 for elimination yields a new mechanism, m_3, constructed as above:

S_1	S_2	S_3	S_4	S_5	S_6		A_1	A_2	A_3	A_4	A_5	A_6
1	1	0	0	0	0	m'_1	0	0	0	−1	1	0
0	0	1	1	0	0	m'_2	0	0	0	0	−1	1
1	1	1	1	1	1	m'_o	0	0	0	0	0	0

Figure 13. The last three rows of the diagonalized setup (shown in its general form in Figure 4) derived by the algorithm of Happel and Sellers (1982, 1983, 1989) for the chemical system of Figure 12. There are two linearly independent overall reactions and one cyclic mechanism.

s_1	s_2	s_3	s_4	s_5	s_6	mechanism	reaction	A_4	A_5	A_6
0	0	−1	−1	−1	−1	m_1	r_1	−1	1	0
0	0	1	1	0	0	m_1	r_2	0	−1	1
0	0	1−2θ	1−2θ	−θ	−θ	m_1	$r=\theta r_1+(1-\theta)r_2$	−θ	2θ−1	1−θ
1	1	0	0	0	0	m_2	r_1	−1	1	0
−1	−1	0	0	−1	−1	m_2	r_2	0	−1	1
1−2θ	1−2θ	0	0	−θ	−θ	m_2	$r=\theta r_1+(1-\theta)r_2$	−θ	2θ−1	1−θ
1	1	0	0	0	0	m_3	r_1	−1	1	0
0	0	1	1	0	0	m_3	r_2	0	−1	1
θ	θ	1−θ	1−θ	0	0	m_3	$r=\theta r_1+(1-\theta)r_2$	−θ	2θ−1	1−θ

Figure 14. The direct mechanisms derived from Figure 4 by the algorithm of Happel and Sellers (1982, 1983, 1989) for the chemical system of Figure 12. Each of the three mechanisms is shown for the linearly independent overall reactions (r_1 and r_2) as well as the general overall reaction r.

- Combining m'_1 and m'_o gives $m'_1+0m'_o=m'_1$ $(=s_1+s_2)$ as the form of the direct mechanism m_2 for r_1.
- Combining m'_2 and m'_o gives $m'_2-m'_o$ as the form of the direct mechanism m_2 for r_2. The mechanism can be written as $-s_1-s_2-s_5-s_6$.

The combination of the two mechanisms for r_1 and r_2 produces the mechanism m_1 for the general reaction $r=\theta r_1+\phi r_2$:

$$m_2:\ \theta(-s_1-s_2-s_5-s_6)+(1-\theta)(s_1+s_2) \ = \ (1-2\theta)s_1+(1-2\theta)s_2-\theta s_5-\theta s_6$$

which is the same as the mechanism that is constructed if s_4 is selected.

Finally, selection of either s_5 or s_6 gives another direct mechanism, m_3, whose form for the overall reaction r is:

$$m_3:\ \theta s_1+\theta s_2+(1-\theta)s_3+(1-\theta)s_4$$

Thus, using r_1 and r_2 as the basis, the algorithm has correctly identified three direct mechanisms, which are shown in Figure 2.14.

The detailed explanation given earlier on the significance of mechanism m_1 also applies to m_2 and m_3. In their forms applicable to the overall reaction r (lines 6 and 9 in Figure 2.14) the mechanisms are quite important, but their submechanisms applicable to r_1 and r_2 (lines 4, 5, 7, and 8 in Figure 2.14) are only means for relating the direct overall mechanisms for the general reaction to the (non-unique) basis set of reactions. It is important to note that the set of direct mechanisms for the general overall reaction is unique, but the *exact expression* derived is dependent on the choice of basis reactions. A form of a direct mechanism for one basis can be transformed to its form for another basis, through a transformation of the set of parameters (in this case, the parameter θ).

The general (indirect) overall mechanism m can be written as a linear combination of the mechanisms m_1 to m_3:

$$m=\mu_1 m_1+\mu_2 m_2+\mu_3 m_3$$

In terms of steps:

$$m = \mu_1\,[(1-2\theta)s_3+(1-2\theta)s_4-\theta s_5-\theta s_6] + \mu_2\,[(1-2\theta)s_1+(1-2\theta)_2-\theta s_5-\theta s_6] + \\ + \mu_3\,[\theta s_1+\theta s_2+(1-\theta)s_3+(1-\theta)s_4]$$

$$\Rightarrow\ m = [\mu_2(1-2\theta)+\mu_3\theta]\ s_1 + [\mu_2(1-2\theta)+\mu_3\theta]\ s_2 + [\mu_1(1-2\theta)+\mu_3(1-\theta)]\ s_3 + \\ + [\mu_1(1-2\theta)+\mu_3(1-\theta)]\ s_4 + (-\theta\mu_1-\theta\mu_2)\ s_5 + (-\theta\mu_1-\theta\mu_2)\ s_6$$

$$\Rightarrow\ m = [\theta(\mu_3-2\mu_2)+\mu_2]\ s_1 + [\theta(\mu_3-2\mu_2)+\mu_2]\ s_2 + [\mu_1+\mu_3-\theta(2\mu_2+\mu_3)]\ s_3 + \\ + [\mu_1+\mu_3-\theta(2\mu_2+\mu_3)]\ s_4 -\theta\ (\mu_1+\mu_2)\ s_5 -\theta\ (\mu_1+\mu_2)\ s_6$$

One may require the coefficients μ_k to satisfy the relationship $\mu_1+\mu_2+\mu_3=1$, so that the net reaction accomplished is r (listed above) rather than a multiple of r. Substitution of $\mu_3=1-\mu_1-\mu_2$ would then eliminate μ_3 from the expression of m.

The expression for the general (overall) mechanism m derived here is quite cumbersome, unlike the much simpler expression relating m and the basis mechanisms m'_1, m'_2, and m'_o (but the latter basis set is not unique). Furthermore, it appears that the expression for m (Equation ?) gives the impression that there are 3 independent parameters (μ_1, μ_2, and θ; assuming the substitution $\mu_3=1-\mu_1-\mu_2$ is first carried out), while there are (as Equation ? shows) only two independent parameters in the expression for the general mechanism. This means that, in expressing the general (indirect) mechanism m as a combination of direct mechanisms for the general reaction, we should only use two of the three direct mechanisms; inclusion of the third is redundant. This should not be surprising in light of the clear explanation given by Happel and Sellers (1982, 1983): Direct mechanisms are *not* necessarily linearly independent.

The Proposed Algorithm. The application of the algorithm proposed in this report is quite easy. There are three intermediate species in this chemical system, and each participates initially in only two steps (or one-step partial mechanisms); thus, elimination of any one intermediate entails construction of one mechanism and removal of two existing (partial) ones. In fact, the three intermediates can be eliminated in parallel, since their steps are disjoint. The resulting setup is shown in Figure 2.15.

The three identified mechanisms (h_1, h_2, and h_3 in Figure 2.15) involve three different direct reactions (r_1, r_2, and r_3, listed in the initial description of the chemical system).

The general (indirect) mechanism for the overall reaction can be viewed as a linear combination of these three mechanisms:

$$h = \chi_1 h_1 + \chi_2 h_2 + \chi_3 h_3$$

In terms of steps, the general mechanism can be written as:

$$h = \chi_1 s_1 + \chi_1 s_2 + \chi_2 s_3 + \chi_2 s_4 + \chi_3 s_5 + \chi_3 s_6$$

The corresponding general overall reaction is:

$$r = (\chi_3 - \chi_1) A_4 + (\chi_1 - \chi_2) A_5 + (\chi_2 - \chi_3) s_6$$

Multiplying the coefficients χ_k by a constant merely multiplies the reaction r and the mechanism h by the same constant. Thus, one can normalize by requiring that $\chi_1+\chi_2+\chi_3=1$, and substituting $\chi_3=1-\chi_1-\chi_2$ to eliminate χ_3.

All three mechanisms of Figure 2.15 are clearly direct and they accomplish three different reactions. Furthermore, the general (indirect) mechanism h can be expressed as a simple combination of direct mechanisms, with the correct number of parameters (χ_1 and χ_2). Thus, there is a close and simple relationship between the direct mechanisms constructed and the general indirect mechanism. The general mechanism also gives an expression for the general reaction. The direct mechanisms derived by the algorithm and the expressions for the general mechanism and the general reaction are unique in every respect. However, the direct mechanisms, the general mechanism, and the general reaction are not related to a basis set of reactions; in fact no basis set of reactions is identified by this procedure.

Analysis. This example points to significant differences not only in the operation of the two methods but even in their definition of direct mechanisms. The differences, already identified in the description of the operation of each algorithm on this example, are summarized here.

Happel and Sellers define a direct mechanism in terms of the particular reaction it accomplishes. In order to construct direct mechanisms for a system with a multiple overall reaction, Happel and Sellers first determine a basis set of linearly independent single reactions. Then, they construct each direct mechanism in terms of its submechanisms for each of the basis reactions; the corresponding direct mechanism for the general reaction is obtained by combining its submechanisms, in the same way the overall reaction can be obtained as a parametric linear combination of the basis reactions. This approach has the advantage that the general form of the overall reaction (in terms of basis reactions) is not only clear but also closely related to the form of the direct mechanisms for the general reaction – although direct mechanisms may be cumbersome because they involve degrees of freedom. However, the relationship between the general (indirect) mechanism and the direct mechanisms is not clear, and does not indicate the number of degrees of freedom in the indirect mechanism. Also, although the set of direct mechanisms is unique, their *exact expressions* are dependent on the choice of basis reactions; a change of basis causes a transformation of the set of parameters in a direct mechanism.

The alternative algorithm proposed here considers a mechanism m_k as direct, regardless of the reaction which it accomplishes, if no other *overall* mechanism uses a subset of the steps used by m_k. Consider mechanism m_1 in Figure 2.14, which shows the direct mechanism according to the Happel and Sellers approach. In the approach presented in this report, only the version of m_1 that accomplishes reaction r_2 is actually direct; the form of m_1 that accomplishes r_1, as well as the form of m_1 for the general reaction (if $\theta \neq 0$, $\theta \neq \frac{1}{2}$) are not direct. Only the mechanisms of Figure 2.15 are direct in our approach. The advantage

s_1	s_2	s_3	s_4	s_5	s_6	mechanism	reaction	A_4	A_5	A_6
1	1	0	0	0	0	h_1	r_1	−1	1	0
0	0	1	1	0	0	h_2	r_2	0	−1	1
0	0	0	0	1	1	h_3	r_3	1	0	−1

Figure 15. The direct mechanisms derived by the algorithm presented in this article, for the chemical system of Figure 12. Three direct mechanisms are shown for three different direct overall reactions. The three mechanisms are linearly independent, but the three reactions are not.

of this view is that direct mechanisms are simpler (each direct mechanism is constructed for some overall reaction which it matches naturally) and involve no degrees of freedom; comparison of Figures 2.14 and 2.15 shows that the algorithm proposed in this report constructed simpler direct mechanisms. For this example, the general indirect mechanism is expressed as a combination of simple direct mechanisms, with the correct number of degrees of freedom. The question of whether this is always true will be addressed in the discussion.

Arbitrary-size Extension of the Chemical System. Figure 2.16 provides an extension of the chemical system considered here, which makes the contrast between the two approaches clearer. There are S=2I steps and A=2I species in all; A_1 through A_I are intermediates, while A_{I+1} through A_{2I} are terminal species. For this extended example, each of the two algorithms will operate in the same way as on the simple chemical system of Figure 2.12, but the differences between the two algorithms are amplified, as will be shown here.

There are I–1 linearly independent reactions. One may choose all reactions of the form

r_i: $A_{I+i} \leftrightarrows A_{I+i+1}$

Specifically, r_1: $A_{I+1} \leftrightarrows A_{I+2}$, r_2: $A_{I+2} \leftrightarrows A_{I+3}$, ..., to r_{I-1}: $A_{2I-1} \leftrightarrows A_{2I}$, form a basis, but the Happel and Sellers algorithm may choose any of a number of possible bases. The general overall reaction can thus be written, as a linear combination of r_1 to r_{I-1}:

$$r = \sum_{i=1}^{I-1} \theta_i r_i$$

with only I–2 of the parameters θ_i considered independent, if the restriction $\sum_{i=1}^{I-1} \theta_i = 1$ is imposed. The Happel and Sellers algorithm will construct one mechanism for each of the basis reactions, and (in this case) one cyclic mechanism (which is unique), $\sum_{i=1}^{2I} s_i$.

In considering each of the 2I steps, the algorithm will find I cases that lead to distinct direct mechanisms. In each of the I cases, a direct submechanism must be constructed for each of the I basis reactions. Then, the direct mechanism for the overall reaction is constructed as a linear combination using the θ_i parameters. In effect, each of the I direct mechanism contains I–1 parameters θ_i (I–2 of which are independent), and is constructed as a combination of I–1 submechanisms. In all, a figure analogous to Figure 2.15 would contain, for this system, I^2 rows, i.e., I^2–I submechanisms for basis reactions and I direct mechanisms for the general reaction. The exact expression for the direct mechanisms derived by the algorithm depends on the particular choice of basis. An expression for the general indirect mechanism as a combination of direct mechanisms would involve I additional parameters μ_i, i.e., 2I–1 parameters in total (of which only 2I–3 are independent).

Application of the algorithm proposed in this report would entail very easy successive elimination of the intermediates (with only one new mechanism constructed for each intermediate; in fact, all intermediates can be eliminated in parallel). The unique resulting set of I direct mechanisms have the form

$h_i = s_{2i-1} + s_{2i}$ (for i=1 to I)

or $A_{I+i} \leftrightarrows A_i \leftrightarrows A_{I+i+1}$ (for i=1 to I–1) and $A_{2I} \leftrightarrows A_I \leftrightarrows A_{I+1}$ (for i=I). Each of these direct mechanisms contains only two steps and accomplishes a different reaction.

$$\sum_{i=1}^{I-1} \theta_i = 1$$

The general indirect mechanism for the overall reaction can be viewed as a combination direct mechanisms:

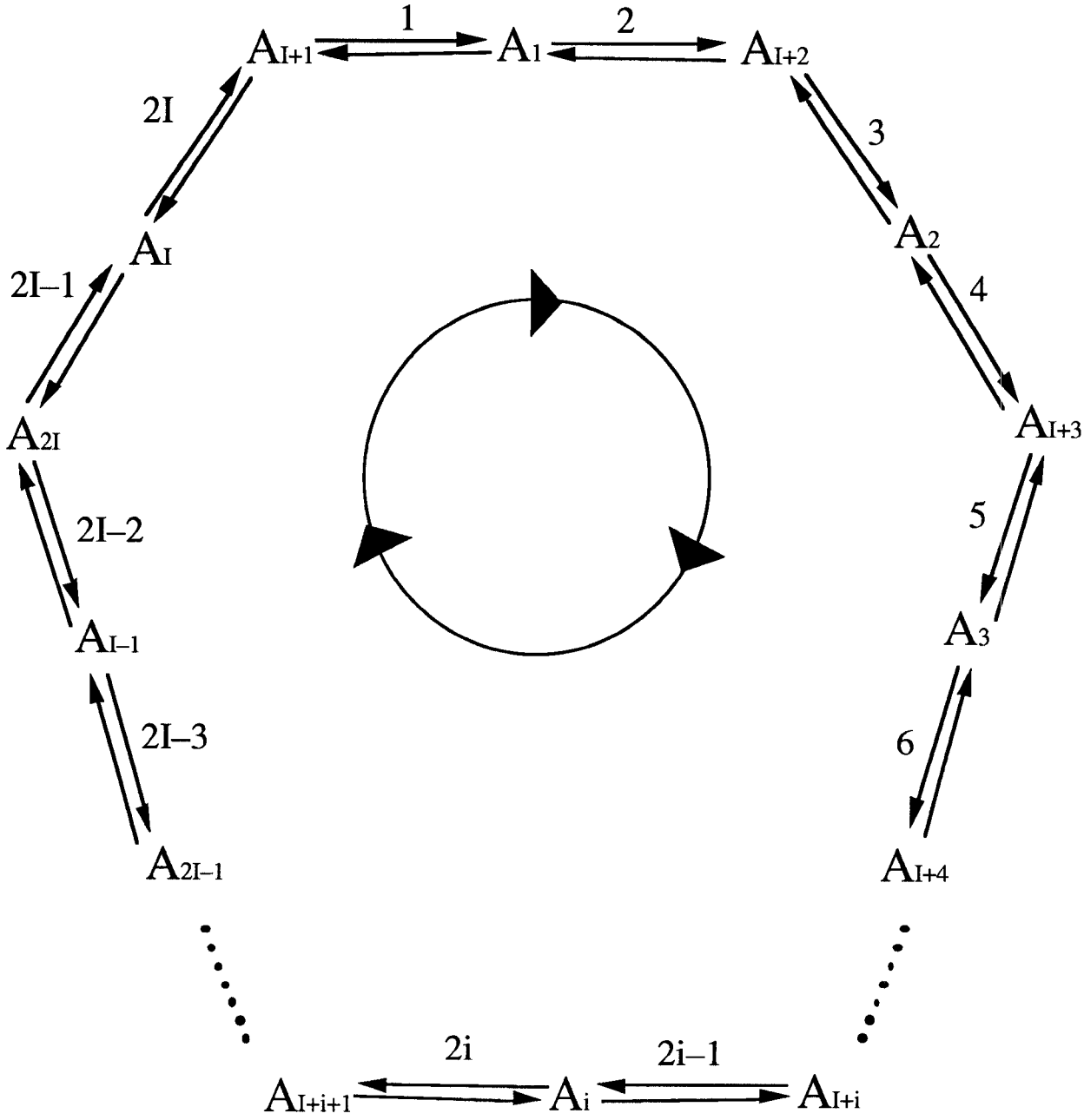


Figure 2.16. An arbitrary-size extension of the chemical system of Figure 2.12. The species A_1 through A_I are intermediates; A_{I+1} through A_{2I} are terminal species. There are 2I steps, numbered in the figure.

$$h = \sum_{i=1}^{I} \chi_i h_i = \sum_{i=1}^{I} \chi_i (s_{2i-1} + s_{2i})$$

and the corresponding general overall reaction as:

$$r = (\chi_I - \chi_1) A_{I+1} + \sum_{i=2}^{I} (\chi_{i-1} - \chi_i) A_{I+i}$$

Normalization may be imposed by requiring that $\sum_{i=1}^{I} \chi_i = 1$.

Thus, simple and unique expressions are derived by the algorithm for the direct mechanisms and the general indirect mechanism.

DISCUSSION

An alternative framework was presented here for the synthesis of direct reaction mechanisms for chemical systems with known species and steps. The basic concepts of direct mechanisms are due to Milner (1964), Happel and Sellers (1982, 1983, 1989), Happel (1986), Sellers (1984, 1989), and Happel *et al.* (1990). Happel and Sellers developed both the chemical insights associated with direct mechanisms and an elegant mathematical formulation of the problem, and their notation was utilized extensively in this report.

The algorithm described here is conceptually based on a method for the synthesis of biochemical pathways (Mavrovouniotis, 1989, Mavrovouniotis *et al.*, 1990), with some significant differences (given in Part 1) in the formulation of the problem, the mathematical treatment, the handling of bidirectional steps, and the elimination of redundant mechanisms (pathways); the last issue was not addressed at all by Mavrovouniotis (1989) and Mavrovouniotis *et al.* (1990), because redundant pathways were rarely encountered in the domain of biochemical systems.

The conceptual framework presented by this report deviates from the concepts established by Happel and Sellers in a subtle aspect of the definition of direct mechanisms. While Happel and Sellers judge directness by comparing a mechanism to other mechanisms accomplishing the same overall reaction, in this report a mechanism must be direct in comparison to all other mechanisms *accomplishing overall reactions* in the chemical system at hand. As can be observed within a multiple-reaction hypothetical system studied in this report, direct mechanisms in the proposed approach do not involve any degrees of freedom.

Happel and Sellers identify direct mechanism for the general form of the overall reaction; for systems with multiple reactions, the direct mechanisms will contain degrees of freedom. They are constructed by identifying first a basis set of reactions (any of the possible basis sets being acceptable). Each direct mechanism is identified as a linear combination of direct submechanisms, one submechanism for each basis reaction; if there are D direct mechanisms and R basis reactions then D×C submechanisms are constructed in all. If one relates the general indirect mechanism to direct mechanisms then only some of the direct mechanisms (forming a basis) should be used. With respect to computational effort, it is noted that each of $\binom{S}{C}$ combinations of C steps must be considered by the algorithm; some (but generally not all) combinations necessitate inversion of a C×C matrix (or determination that the matrix is singular).

The algorithm presented in this report does not consider linear independence. It constructs direct mechanisms which do not contain degrees of freedom, without focusing on particular reactions. In a simple example with a multiple reaction considered in Figures 2.12 to 2.16, the general indirect mechanism could be expressed as a linear combination of the direct mechanisms, with the correct degrees of freedom. To maintain the correct degrees of freedom in the general case, however, some restrictions must be used on the way in which direct mechanisms are allowed to be combined. Certain restrictions can be derived from the rigorous mathematical framework of Sellers (1984, 1989), and in particular Theorem

2 in Sellers (1989, p. 299). Specifically, in order to express an indirect mechanism as a linear combination of direct mechanisms, the direct mechanisms must include only those steps used by the indirect mechanism and *in the same direction*; in other words, one may not combine two (or more) direct mechanisms if the combination leads to utilization of the same step in both directions (regardless of whether the step cancels out altogether or not). In comparing this statement to Theorem 2 of Sellers (1989) one should bear in mind the difference in the definition of direct mechanisms. One should also note that the definition of direct mechanisms proposed in this report actually does not preclude certain cyclic mechanisms! In chemical systems that contain only unidirectional steps the condition given above holds for any combination of direct mechanisms (with positive combination coefficients). In the bidirectional example considered in this report, the condition also happens to hold for all combinations. In general, however, this condition *along with additional requirements* must be imposed allow only some of the possible combinations of direct mechanisms. The analysis of the requirements and their algorithmic ramifications will be considered in a forthcoming report. The requirements actually lead to a more sophisticated version of the algorithm which also avoids redundant mechanisms altogether (rather than recognizing and discarding them).

Finally, a major difference in the operation of the two algorithms arises in their treatment of directionality restrictions on mechanism steps. Directionality is ignored by Happel and Sellers until the final mechanisms are constructed, at which point infeasible mechanisms can be rejected. Directionality is sometimes especially cumbersome in the Happel and Sellers algorithm. For example, in the methanol example of Happel *et al.* (1990), the direct submechanisms for one of the basis reactions violate directionality restrictions. To address this deficiency, the basis reaction is replaced by another; it is not clear whether this difficulty must be treated manually whenever it is encountered or addressed systematically through an extra phase in the Happel and Sellers algorithm. The algorithm presented here can take into account any restrictions on permissible directions of steps as the mechanisms are constructed. This may lead to significant gains in efficiency, for chemical systems that contain many unidirectional steps.

REFERENCES

Happel, J. *Isotopic Assessment of Heterogeneous Catalysis*, Academic Press, **1986**.

Happel, J., and Sellers, P. H. "Analysis of the Possible Mechanisms for a Catalytic Reaction System" *Advances in Catalysis.*, **1983**, *32*, 273-323.

Happel, J., and Sellers, P. H. "Multiple Reaction Mechanisms in Catalysis" *Ind. Eng. Chem. Fund.*, **1982**, *21*, 67-76.

Happel, J., and Sellers, P. H. "The Characterization of Complex Systems of Chemical Reactions" *Chem. Eng. Comm.*, **1989**, *83*, 221-240.

Happel, J., Sellers, P. H., and Otarod, M. "Mechanistic Study of Chemical Reaction Systems" *Ind. and Eng. Chem. Res.*, **1990**, *29*, 1057-1064.

Horiuti, J. "Theory of Reaction Rates as based on the Stoichiometric Number Concept" *Ann. N.Y. Acad. Sci.*, **1973**, *213*, 5-30.

Horiuti, J., and Nakamura, T. "On the Theory of Heterogeneous Catalysis" *Advances in Catalysis*, **1967**, *17*, 1-74.

Lee, L.-S., and Sinanoğlu, O. "Reaction Mechanisms and Chemical Networks — Types of Elementary Steps and Generation of Laminar Mechanisms" *Zeitschrift für Physikalische Chemie*, **1981**, *124*, 129-160.

Mavrovouniotis, M. L. "Computer-Aided Design of Biochemical Pathways" Ph.D. Thesis, Massachusetts Institute of Technology, **1989**.

Mavrovouniotis, M. L. "Synthesis of Reaction Mechanisms Consisting of reversible and irreversible steps: 2. Formalization and Analysis of the Synthesis Algorithm" *Ind. and Eng. Chem. Res.*, **1991**, this issue.

Mavrovouniotis, M. L., Stephanopoulos, G., and Stephanopoulos, G. "Computer-Aided Synthesis of Biochemical Pathways" *Biotechnology and Bioengineering*, **1990**, *36*, 1119-1132.

Milner, P. C. "The Possible Mechanisms of Complex Reactions Involving Consecutive Steps" *J. Electrochem. Soc.*, **1964**, *111*, 228-232.

Sellers, P. H. "Combinatorial Aspects of Enzyme Kinetics", in *Applications of Combinatorics and Graph Theory in the Biological and Social Sciences*, F. Roberts (ed.), Springer-Verlag, **1989**.

Sellers, P. H. "Combinatorial Classification of Chemical Mechanisms" *SIAM J. Appl. Math.*, **1984**, *44*, 784-792.

Sinanoğlu, O. "Theory of Chemical Reaction Networks. All Possible Mechanisms or Synthetic Pathways with Given Number of Reaction Steps or Species" *J. Am. Chem. Soc.*, **1975**, *97*, 2309-2320.

Sinanoğlu, O., and Lee, L.-S. "Finding All Possible *a priori* Mechanisms for a Given Type of Overall Reaction" *Theoret. Chim. Acta (Berl.)*, **1978**, *48*, 287-299.

Temkin, M. I. "The Kinetics of Some Industrial Heterogeneous Catalytic Reactions" *Advances in Catalysis*, **1979**, *26*, 173-291.

Temkin, M. I. "The Kinetics of Transfer of Labeled Atoms by Reaction" *Ann. N.Y. Acad. Sci.*, **1973**, *213*, 79-89.

Yarlagadda, P.S., Morton, L.A., Hunter, N.R., and Gesser, H.D. "Direct Conversion of Methane to Methanol in a Flow Reactor" *Ind. Eng. Chem. Res.*, **1988**, *27*, 252-256.