Construction of Chinese-English Semantic Hierarchy for Information Retrieval

Gina-Anne Levow
Institute for Advanced Computer Studies
University of Maryland

Bonnie Dorr
Institute for Advanced Computer Studies
and
Department of Computer Science
University of Maryland

Dekang Lin
Department of Computer Science
University of Manitoba

E-mail: gina,bonnie,lindek@umiacs.umd.edu

Abstract

This paper describes an approach to large-scale construction of a semantic hierarchy for Chinese verbs. Leveraging off of an existing Chinese conceptual database called HowNet and a Levin-based English verb classification, we use thematic-role information to create links between Chinese concepts and English classes. The resulting hierarchy is used for multilingual lexicons in an English-Chinese cross-language information retrieval application. We demonstrate a structured syntax interface that exploits this large-scale hierarchy and its linkages to WordNet for English-Chinese cross-language information retrieval.

1 Introduction

The growing quantity of online multilingual information has created an urgent need for rapid construction of lexical resources. Automatic and semi-automatic techniques for lexical acquisition are more critical now than ever before as it becomes infeasible to produce adequate semantic representations on a large scale by human labor alone.

We describe an approach to large-scale construction of a semantic hierarchy for Chinese verbs. Leveraging off of an existing classification of English verbs called EVCA (English Verbs Classes and Alternations) [12] and a Chinese conceptual database called HowNet [25, 24, 23] (http://www.how-net.com), we use thematic-role information (e.g., a mapping between the HowNet "Patient" and the EVCA-based "Th(eme)") to create links between Chinese concepts and English classes. Each Chinese-English link is additionally associated with a sense from WordNet [13], thus producing a new Asian companion to the current (Euro)WordNet initiative. Finally, the EVCA semantic class, thematic role mapping, and a canonical English word are used to produce a full lexical conceptual structure (LCS) entry for the verb.

We use the resulting lexicons to determine word senses in a cross-language information retrieval application, where the degree of accuracy is significantly improved over the weak alternative of a bilingual word list.

Finally, we will describe a Chinese-English cross-language information retrieval system that exploits this lexicon to improve word sense disambiguation. The system uses a structured syntax interface to facilitate mapping from the surface form of the user's query to a semantically rich interlingual representation. Furthermore, it relies on structural matching of thematic roles and taxonomic similarity measures using linkages to WordNet that are derived as part of the construction of the semantic hierarchy.

2 Constructing Rich Cross-language Lexical Resources

Ordinary within-language lexical ambiguity is exacerbated in the cross-language context, as each sense of a word may have many alternate translations. For example, the Chinese verb 拉 (*la* corresponds to a wide range of English glosses—even if we examine only the verb translations—in the Optilex[1] Chinese-English dictionary: *slash*, *cut*, *chat*, *pull*, *drag*, *transport*, *move*, *raise*, *help*, *implicate*, *involve*, *defecate*, *pressgang*.[2] Our work provides a framework for disambiguating such cases in a given context by associating certain of these senses (e.g., *transport*, *move*) with one HowNet concept (e.g., |Transport|) while associating other senses (e.g., *help*) to another HowNet concept (e.g., |help|).

2.1 Related Work: Mapping across Semantic Hierarchies

Several researchers have investigated the problem of assigning class-based senses to verbs [2], [7], [6] [5], [11], [14] [18], [20], [19], and [22]. This work extends the techniques described by [20], which used a concept space to produce a hierarchical organization of Chinese verbs. The extensions include the use of the entire EVCA database rather than a small set of verbs (the *break* class) and the provision of a thematic-role based filter. We adopt a technique that is similar in flavor to the intersective-class approach of [2], with the following extensions: (1) Concept alignment across two different language hierarchies (Chinese and English) rather than one; (2) Mappings between Chinese and English thematic roles; and (3) Hooks into WordNet senses for both languages.

The EVCA classes used in this work include 485 total classes, each hand-tagged with WordNet senses and thematic-role specifications. Mapping English roles to their Chinese counterparts is the primary aid in associating WordNet senses with Chinese verbs; the thematic-role mappings are used as a guideline for selecting the appropriate entry in EVCA, which in turn is associated with a WordNet sense. The Chinese concept hierarchy HowNet is an on-line conceptual commonsense knowledge base that contains hierarchical information relating concepts as well as a thematic-role specification to the associated Chinese word within the verb hierarchy which is our focus.

[1] Optilex is the machine-readable version of the CETA dictionary, licensed from the MRM corporation, Kensington, MD.

[2] Optilex is a large (600k entries) machine readable Chinese-English dictionary; although this dictionary is in some ways exhaustive, there is no encoding of part-of-speech information, but see [18] for a description of a procedure that extracts verbs automatically from Optilex.

2.2 Mapping Between Chinese HowNet and English EVCA

The mapping between Chinese HowNet and English EVCA involves three steps, illustrated in Figure 1:

(1) Produce all possible English Optilex *glosses* (translations) for all 12342 Chinese verbs in HowNet and associate each Chinese verb with one or more of the HowNet concepts. [HowNet Class+Word+Gloss Figure 1]

(2) Associate each verb-to-concept candidate with one or more of the 485 EVCA classes—forming an average of 2 thousand verb-to-class entries per HowNet concept (on the order of 1 million verb-to-class candidates, total). [EVCA Class Mapping Figure 1]

(3) For each HowNet concept, partition the associated Chinese-English pairs into groups whose English glosses correspond EVCA classes. This requires three steps:

   a. Order the candidate EVCA classes so that the highest-ranking classes are those that contain the highest number of English verbs matching the Optilex glosses. [Ranking by EVCA Class: Figure 1]

   b. In cases where a tie-breaker is needed, reorder the candidate EVCA classes according to the degree to which the thematic-role specification in HowNet concept matches that of EVCA class. [Ranking by Thematic Role Mapping: FIgure 1]

   c. For each Chinese-English entry associated with the HowNet concept, assign the highest ranking candidate EVCA class. [Output Mapping: Figure 1]

The process of associating EVCA classes with Chinese verbs relies on a massive filtering of spurious class assignments. For example, the |Establish| HowNet concept is ultimately associated with only two EVCA classes, 29.2.c and 26.4.a (Characterize and Create), but it initially had 29 potential EVCA class assignments. One example of an EVCA class that was ruled out is the Change of State class, 45.4.a, associated with the Optilex translation *colonize* for the Chinese verb 殖民 (*zhimin*). Although this is a perfectly valid EVCA class assignment for the HowNet concept |Colonize|, it is not appropriate for the |Establish| HowNet concept. Because this class is ranked 8th for |Establish|—as opposed to 1st and 2nd place ranking

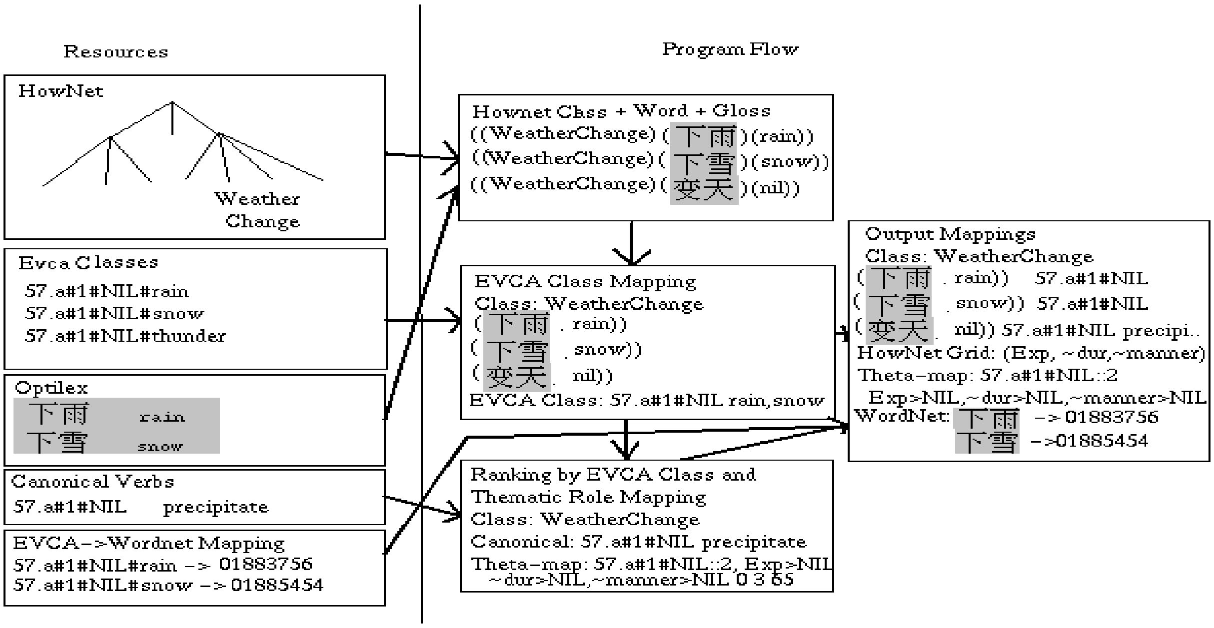


Figure 1. Resources and Processing Stages for Mapping Chinese HowNet and English EVCA, including linkages to English WordNet

for 29.2.c and 26.4.a, respectively—this assignment is ruled out by our algorithm.

3 Building a Chinese Lexicon with Lexical Conceptual Structure Entries and WordNet Links

The technique described above creates a bridge between entries in the Chinese HowNet conceptual hierarchy and the EVCA semantic classes. Next we demonstrate how these thematic role and semantic class mappings are combined to produce a rich lexical resource for cross-language information retrieval, with a focus on use of event structure for word sense disambiguation.

3.1 Lexical Conceptual Structure

Lexical Conceptual Structure is a language-independent representation used in the NLP component of an implemented foreign language tutoring system [5] and an interlingual machine translation system [18]. Our goal is to examine the use of this representation in the context of cross-language information retrieval. First, we show how an LCS-based classification can be used to develop a cross-language lexical acquisition approach that contributes both toward the enrichment of existing online resources (the HowNet semantic hierarchy and the Levin-based verb semantic classification system) and toward the development of lexicons containing more complete information than is provided in any of these resources alone. Next, we demonstrate the applicability of LCS to the problem of cross-language information retrieval.

3.1.1 Components of Lexical Conceptual Structure

One of the types of knowledge that must be captured in cross-language information retrieval is linguistic knowledge at the level of the lexicon, which covers a wide range of information types, such as verbal subcategorization for events (e.g., that a transitive verb such as "hit" occurs with an object noun phrase), featural information e.g., that the direct object of a verb such as "frighten" is animate), thematic information (e.g., that "John" is the agent in "John hit the ball"), and lexical-semantic information (e.g., that spatial verbs such as "throw" are conceptually distinct from verbs of possession such as "give"). By modularizing the lexicon, we treat each information type separately, thus allowing us to vary the degree of dependence on each level, so that we can address the question of how much knowledge is necessary for the success of the particular NLP application.

The most intricate component of lexical knowledge is the lexical-semantic information, which is encoded in the form of Lexical Conceptual Structure (LCS) as formulated by Dorr [3, 4] based on work by Jackendoff [9, 10]. The LCS approach views semantic representation as a subset of conceptual structure, the language of mental representation, as in [9, 10]. This approach includes *types* such as Event and State, which are specialized into *primitives* such as GO, STAY, BE, GO-EXT, and ORIENT. We add a manner component $[_{\text{Manner}}$ JOGGINGLY$]$ to distinguish among verbs, e.g. *run*, *walk*, and *jog*. The full representation for *John jogged to school* is therefore the representation below, roughly 'John went to the school by jogging':

$$
\begin{array}{l}
[_{\text{Event}}\ \text{GO}_{\text{Loc}} \\
\quad ([_{\text{Thing}}\ \text{JOHN}], \\
\quad\ [_{\text{Path}}\ \text{TO}_{\text{Loc}} \\
\quad\quad ([_{\text{Thing}}\ \text{JOHN}], \\
\quad\quad\ [_{\text{Position}}\ \text{AT}_{\text{Loc}}\ ([_{\text{Thing}}\ \text{JOHN}], [_{\text{Thing}}\ \text{SCHOOL}])])], \\
\quad\ [_{\text{Manner}}\ \text{JOGGINGLY}])]
\end{array}
$$

3.1.2 Acquisition of an LCS lexicon

As described in [5], we use Levin's publicly available online index [12] as a starting point for building LCS-based verb entries. (We have enhanced this database to include approximately 3,000 additional verbs, for a total of 10,000 verb entries.) While this index provides a unique and extensive catalog of verb classes, it does not define the underlying meaning components of each class. One of the main contributions of our work is that it provides a relation between Levin's classes and meaning components as defined in the LCS representation.

Three inputs are required for acquisition of verb entries: a semantic class, a thematic grid, and a canonical English verb. The output is a Lisp-like expression corresponding to the LCS representation. Given that we have mapped the HowNet grid entries to LCS-based grid entries, we are able to produce the LCS's for Chinese in the same way that we produce entries for English.

Below we present the case of generating an LCS entry for the Chinese verb 接触 (*to touch*). The input/output for our acquisition procedure is shown here:

Acquisition of LCS for:
Chinese: 接触
English: *touch*
Input: 47.8.f; `_th_loc`; "touch"
Output:

```
(be loc (* thing 2)
        (at loc (thing 2) (* thing 11))
        (touchingly 26))
```

4 Application to Chinese-English Cross-language Information Retrieval

We apply this expanded semantic hierarchy to the development of an interactive information retrieval system employing Lexical Conceptual Structure Query Translation (LQT).

4.1 Related Work: Translation in Cross-Language Information Retrieval

A common approach to transforming documents and queries in different languages into a common indexing space for cross-language information retrieval (CLIR) is to translate either the document or the queries into a single language [17]. Due to the time and computational expense of translation, query translation is often preferred over document translation, although document translation often produces superior results. A prevalent technique for query translation is referred to as dictionary query translation (DQT) in which the system looks up each term in the query in a bilingual dictionary or term list and replaces each term with one or more corresponding document language terms [15]. A variety of methods have been applied to translation term selection to cope with the problem of translation ambiguity, where one source language term translates to more than one target language alternative [16], [1], [8]. These techniques include selecting every query, the first N translations according to some ranking strategy, and those that co-occur with candidate translation of other terms in the query.

4.2 System Design

LQT relies on the use of an interlingual representation, LCS, to translate the user's query into the document language for information retrieval. The LCS encodes deep semantic analysis and subcategorization information. This information facilitates word sense disambiguation in query translation by exploiting the grammatical context of the term. In our current system, we use a structured syntax interface, called MADLIBS (Maryland Action Detection / Language-Independent Browsing and Search), to ensure that the user's query is fully analyzable for application of LQT. Specifically, for each word in the LCS lexicon we produce a simple "composed" LCS for each thematic role structure associated with the word, instantiating each role position with a dummy lexical entry, e.g. "someone-1" or "something-2". We then convert this version of the LCS into a template for user input, by generating a syntactically correct surface sentence realization using the Nitrogen generation system from ISI. We now have a mapping from surface forms to interlingual structures.

To guide user input, we developed the interface illustrated below in Figure 2. The positions in the sentence realization that correspond to the thematic roles appear as boxes for free-form user input. The interface allows querying of either English or Chinese documents; we will focus on the cross-language variant in the remainder of this discussion.

To construct the query, the surface template retrieves its underlying CLCS structure, complete with the input words filling the thematic role positions. This correspondence between surface form and thematic structure performs an initial phase of sense disambiguation, identifying the subset of possible senses with this argument structure. To perform translation of the query, we perform a structural match of the query against a database of LCS structures, built from the thematic hierarchy. Depending on language choice, we consult different databases and return words with corresponding LCS structure. This structural match also directly exploits the thematic role information built into the expanded thematic hierarchy.

The system permits two forms of matching: exact and relaxed, selected with the pull-down item in the interface. Exact match compares both structure and manner constants and relies only on the database selection. Relaxed match performs a second phase of processing after the structural match, employing the WordNet correspondences produced by the thematic hierarchy. This method computes similarity between the original term and the candidate translations returned by the structural match building on Resnik's [21] technique for computing taxonomic similarity. In all cases, the top N scoring candidates are returned.

We currently perform no additional analysis of noun phrases entered in the thematic role position, though a fuller treatment of nominalized events is planned. We instead apply basic DQT techniques, using a lexicon built from the Linguistic Data Consortium's [3] English-Chinese term list augmented with the result of inverting the Optilex lexicon for words with single word translations, for the Chinese document case. Again, we select the top N translation alternatives.

The translation terms identified by structural match, taxonomic match and word-for-word translation form a bag of words that comprise the query to an information retrieval system. We use a version of the SMART information retrieval system, modified for 2-byte encodings of Chinese characters. Results are displayed interactively as well (see Figure 3), in the user's choice of source document language, Systran machine translation, or "gist", a word-for-word translation technique that provides multiple ranked alternate translations (see Figure 4).

5 Summary and Future Work

We have presented an approach to aligning two large-scale online resources, HowNet and EVCA. The lexicon resulting from this approach is large-scale, containing more than 17000 Chinese-English conceptual links. The technique for producing these links involves matching semantic-role specifications in HowNet with those in EVCA. Because each Chinese-English link is additionally associated with a WordNet sense, we see this resource as the first step toward producing a new Asian language companion to ongoing (Euro)WordNet initiatives. We have also described a system which exploits both the lexicon and its connections with EVCA classes and WordNet to improve word sense disambiguation in Chinese-English cross-language information retrieval. We plan to perform a quantitative evaluation of the effectiveness of this form of information retrieval on event-based queries.

We are currently investigating the use of the lexicon for word-sense disambiguation in machine-translation and cross-language information retrieval in conjunction with other established corpus techniques for sense selection such as corpus cooccurrence. As we saw above the Chinese verb 拉 (*la*) has several possible translations, but not all of these will be appropriate in every context. If we can determine which HowNet concept corresponds to 拉 (*la*), then we will translate it appropriately. For example, if the HowNet concept is |Transport|, the translation would be *ship* or *transport*, but not *slash*, *chat*, *implicate*, etc. We can detect which HowNet class is appropriate by examining the other words in the sentence. For this word, co-occurrence with a specific word in argument position is a particularly powerful disambiguating cue. If those words co-occur with *other* Chinese verbs associated with a particular HowNet concept (as determined

[3] www.ldc.upenn.edu

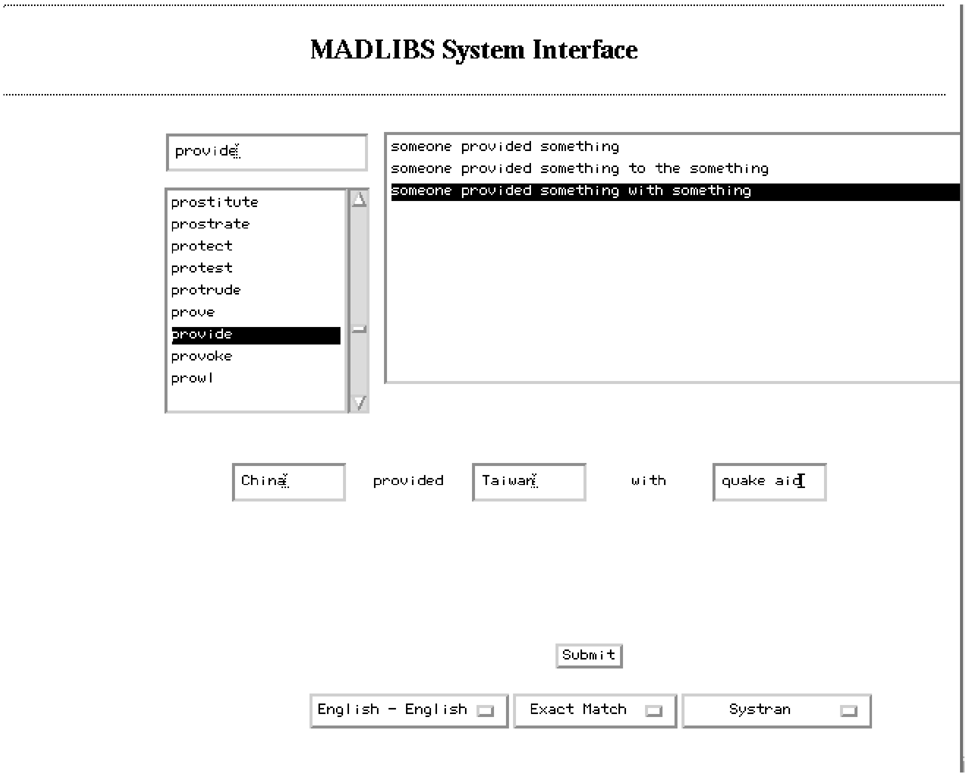
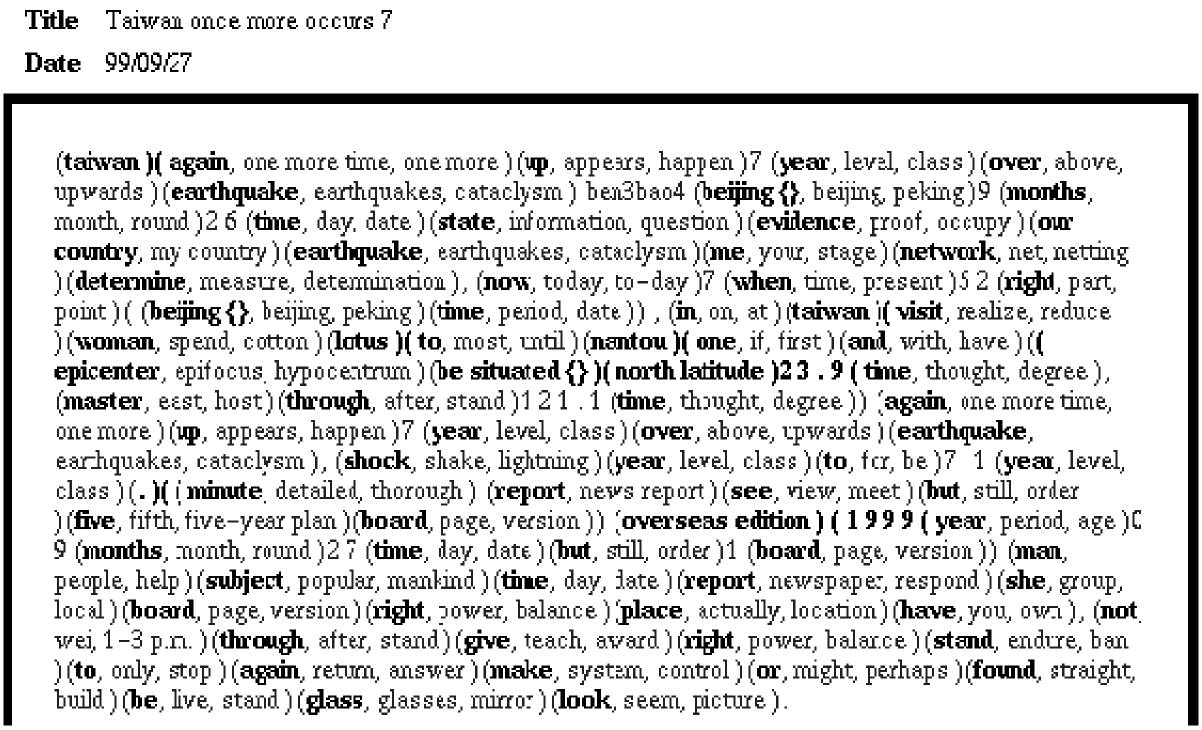


Figure 2. Structured Syntax Input Interface

1	99/10/07	other day, Fuzhou resident	Glossed	Systran MT	Chinese
2	99/09/27	Fujian Province	Glossed	Systran MT	Chinese
3	99/09/27	on September 26th, Taiwan	Glossed	Systran MT	Chinese
4	99/09/27	Taiwan once more occurs 7	Glossed	Systran MT	Chinese
5	99/09/27	Ç¿Őð unceasing typhoon	Glossed	Systran MT	Chinese
6	99/09/27	motherland mainland	Glossed	Systran MT	Chinese
7	99/10/07	compatriot kisses	Glossed	Systran MT	Chinese
8	99/10/04	Red Cross Society of China	Glossed	Systran MT	Chinese
9	99/09/27	each kind of activity	Glossed	Systran MT	Chinese
10	99/07/07	eulogizes Taiwan	Glossed	Systran MT	Chinese

1 2 3 4 5 6 7 8 9 10 11 Next >>>

Figure 3. Selection Interface



Title Taiwan once more occurs 7

Date 99/09/27

(taiwan)(again, one more time, one more)(up, appears, happen)7 (year, level, class)(over, above, upwards)(earthquake, earthquakes, cataclysm) ben3bao4 (beijing {}, beijing, peking)9 (months, month, round)2 6 (time, day, date)(state, information, question)(evidence, proof, occupy)(our country, my country)(earthquake, earthquakes, cataclysm)(me, your, stage)(network, net, netting)(determine, measure, detemination), (now, today, to-day)7 (when, time, present)5 2 (right, part, point)((beijing {}, beijing, peking)(time, period, date)) , (in, on, at)(taiwan)(visit, realize, reduce)(woman, spend, cotton)(lotus)(to, most, until)(nantou)(one, if, first)(and, with, have)((epicenter, epifocus, hypocentrum)(be situated {})(north latitude)2 3 . 9 (time, thought, degree), (master, east, host)(through, after, stand)1 2 1 . 1 (time, thought, degree)) (again, one more time, one more)(up, appears, happen)7 (year, level, class)(over, above, upwards)(earthquake, earhquakes, cataclysm), (shock, shake, lightning)(year, level, class)(to, for, be)7 1 (year, level, class)(.)((minute, detailed, thorough) (report, news report)(see, view, meet)(but, still, order)(five, fifth, five-year plan)(board, page, version)) (overseas edition) (1 9 9 9 (year, period, age)0 9 (months, month, round)2 7 (time, day, date)(but, still, order)1 (board, page, version)) (man, people, help)(subject, popular, mankind)(time, day, date)(report, newspaper, respond)(she, group, local)(board, page, version)(right, power, balance)(place, actually, location)(have, you, own), (not, wei, 1-3 p.m.)(through, after, stand)(give, teach, award)(right, power, balance)(stand, endure, ban)(to, only, stop)(again, return, answer)(make, system, control)(or, might, perhaps)(found, straight, build)(be, live, stand)(glass, glasses, mirror)(look, seem, picture).

Figure 4. Presentation Interface: "Gisted" Format

through a corpus analysis), then it is likely that that HowNet concept is the appropriate one for the Chinese verb. That is, if we find other verbs from a given HowNet concept occurring in the same context, then we can hypothesize that this particular verb has the meaning of this HowNet concept.

Another area of investigation is the use of a WordNet-based distance metric (e.g., the information-content approach of [21]) for additional pruning power in the HowNet-to-EVCA alignment. Because each of the entries in the EVCA classification is associated with a WordNet sense, it is possible to rule out certain class assignments for a given HowNet concept by examining semantic distance between the Optilex glosses for a particular Chinese word and the glosses for other words associated with that concept.

Acknowledgements The University of Maryland authors are supported, in part, by PFF/PECASE Award IRI-9629108, DOD Contract MDA904-96-C-1250, and DARPA/ITO Contract N66001-97-C-8540. Dekang Lin is supported by Natural Sciences and Engineering Research Council of Canada grant OGP121338.

References

[1] Lisa Ballesteros and W. Bruce Croft. Phrasal translation and query expansion techniques for cross-language information retrieval. In *Proceedings of the 20th International ACM SIGIR Conference on Research and Development in Information Retrieval*, July 1997.

[2] Hoa Trang Dang, Karin Kipper, Martha Palmer, and Joseph Rosenzweig. Investigating Regular Sense Extensions Based on Intersective Levin. In *ACL/COLING 98, Proceedings of the 36th Annual Meeting of the Association for Computational Linguistics* (joint with the 17th International Conference on Computational Linguistics), pages 293–306, Montreal, Canada, August 10-14 1998.

[3] Bonnie J. Dorr. *Machine Translation: A View from the Lexicon.* The MIT Press, Cambridge, MA, 1993.

[4] Bonnie J. Dorr. Machine Translation Divergences: A Formal Description and Proposed Solution. *Computational Linguistics*, 20(4):597–633, 1994.

[5] Bonnie J. Dorr. Large-Scale Acquisition of LCS-Based Lexicons for Foreign Language Tutoring. In

Proceedings of the ACL Fifth Conference on Applied Natural Language Processing (ANLP), pages 139–146, Washington, DC, 1997.

[6] Bonnie J. Dorr and Douglas Jones. Acquisition of Semantic Lexicons: Using Word Sense Disambiguation to Improve Precision. In *Proceedings of the Workshop on Breadth and Depth of Semantic Lexicons, 34th Annual Conference of the Association for Computational Linguistics*, pages 42–50, Santa Cruz, CA, 1996.

[7] Bonnie J. Dorr and Douglas Jones. Acquisition of semantic lexicons: Using word sense disambiguation to improve precision. In Evelyne Viegas, editor, *Breadth and Depth of Semantic Lexicons*. Kluwer Academic Publishers, Norwell, MA, 1999.

[8] David A. Hull and Gregory Grefenstette. Experiments in multilingual information retrieval. In *Proceedings of the 19th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 1996. http://www.xerox.fr/people/grenoble/hull/papers/sigir96.ps.

[9] Ray Jackendoff. *Semantics and Cognition*. The MIT Press, Cambridge, MA, 1983.

[10] Ray Jackendoff. *Semantic Structures*. The MIT Press, Cambridge, MA, 1990.

[11] Douglas Jones, Robert Berwick, Franklin Cho, Zeeshan Khan, Karen Kohl, Naoyuki Nomura, Anand Radhakrishnan, Ulrich Sauerland, and Brian Ulicny. Verb Classes and Alternations in Bangla, German, English, and Korean. Technical report, Massachusetts Institute of Technology, 1994.

[12] Beth Levin. *English Verb Classes and Alternations: A Preliminary Investigation*. University of Chicago Press, Chicago, IL, 1993.

[13] George A. Miller and Christiane Fellbaum. Semantic Networks of English. In Beth Levin and Steven Pinker, editors, *Lexical and Conceptual Semantics, Cognition Special Issue*, pages 197–229. Elsevier Science Publishers, B.V., Amsterdam, The Netherlands, 1991.

[14] Naoyuki Nomura, Douglas A. Jones, and Robert C. Berwick. An architecture for a universal lexicon: A case study on shared syntactic information in Japanese, Hindi, Ben Gali, Greek, and English. In *Proceedings of COLING-94*, pages 243–249, 1994.

[15] Douglas W. Oard. Alternative approaches for cross-language text retrieval. In *AAAI Symposium on Cross-Language Text and Speech Retrieval*. American Association for Artificial Intelligence, March 1997. http://www.glue.umd.edu/~oard/research.html.

[16] Douglas W. Oard. A comparative study of query and document translation for cross-language information retrieval. October 1998.

[17] Douglas W. Oard and Bonnie J. Dorr. A survey of multilingual text retrieval. Technical Report UMIACS-TR-96-19, University of Maryland, Institute for Advanced Computer Studies, April 1996. http://www.glue.umd.edu/~oard//research.html.

[18] Mari Broman Olsen, Bonnie J. Dorr, and Scott C. Thomas. Enhancing Automatic Acquisition of Thematic Structure in a Large-Scale Lexicon for Mandarin Chinese. In *Proceedings of the Third Conference of the Association for Machine Translation in the Americas, AMTA-98, in Lecture Notes in Artificial Intelligence, 1529*, pages 41–50, Langhorne, PA, October 28-31 1998.

[19] Martha Palmer and Joseph Rosenzweig. Capturing motion verb generalizations with synchronous tags. In *Proceedings of the Second Conference of the Association for Machine Translation in the Americas*, Montreal, Quebec, Canada, 1996.

[20] Martha Palmer and Zhibao Wu. Verb Semantics for English-Chinese Translation. *Machine Translation*, 10(1–2):59–92, 1995.

[21] Philip Resnik. Using information content to evaluate semantic similarity in a taxonomy. In *Proceedings of IJCAI-95*, pages 448–453, Montreal, Canada, 1995.

[22] Patrick Saint-Dizier. Semantic Verb Classes Based on 'Alternations' and on WordNet-like Semantic Criteria: A Powerful Convergence. In *Proceedings of the Workshop on Predicative Forms in Natural Language and Lexical Knowledge Bases*, pages 62–70, Toulouse, France, 1996.

[23] Dong Zhendong. Enlightment and Challenge of Machine Translation. *Shanghai Journal of Translators for Science and Technology*, 1:9–15, 1988.

[24] Dong Zhendong. Knowledge Description: What, How and Who? In *Proceedings of International Symposium on Electronic Dictionary*, page 18, Tokyo, Japan, 1988.

[25] Dong Zhendong. MT Research in China. In *Proceedings of International Conference on New Directions in Machine Translation*, pages 85–91, Budapest, 1988.