Three-Level Knowledge Representation of Predicate-Argument Mapping for Multilingual Lexicons

Chinatsu Aone and Doug McKee
Systems Research and Applications
2000 15th Street North
Arlington, VA 22201
aonec@sra.com, mckeed@sra.com

lexical levels, automatic acquisition

1 Introduction

Constructing a semantic concept hierarchy covers only part of the problem in building any knowledge-based natural language processing (NLP) system: Language understanding requires mapping from syntactic structures into conceptual (i.e. interlingua) representation (henceforth verb-argument mapping), while language generation requires the inverse mapping. Jacobs [13] proposed a solution in which he augmented his concept hierarchy in the knowledge base with verb-argument mapping information. While his solution works for monolingual understanding/generation, associating each semantic concept of a verb with English specific mapping information in the knowledge base is not ideal for an interlingua-based NLP system, since verb-argument mapping varies greatly among languages.

In this paper, we discuss how the lexicon of our interlingua-based NLP system[1] abstracts the language-dependent portion of predicate-argument mapping information from the core meaning of verb senses (i.e. *semantic concepts* as defined in the knowledge base). We also show how we represent this mapping information in terms of cross-linguistically generalized mapping types called *situation types* and word sense-specific *idiosyncrasies*. This three-level representation enables us to keep the language-independent knowledge base intact and make efficient use of lexical data by inheritance. We take advantage of large multilingual corpora to automatically acquire language- and domain-specific lexical information.

In addition, this representation framework easily deals with some of the MT divergences which arise from predicate-argument mapping mismatches. Namely, verbs in different languages denoting the same semantic concept do not always represent their thematic roles with the same syntactic structures (cf. McCord [17], Dorr [10]). For example, "I like Mary" is translated into "María me gusta a mí", where the subject and the object in the English sentence are the object and the subject in the Spanish sentence. Furthermore, a verb with two different predicate-argument mappings may correspond to two different verbs in another language. For example, the English word "break," when AGENT is expressed as in "John broke the window," is mapped to Japanese word "kowasu", while without AGENT, as in "The window broke," "break" should map to "kowareru".

[1] Our interlingua-based NLP system consists of Solomon, a broad-coverage text understanding system for English, Spanish and Japanese, and an experimental generation system which constructs sentences from interlingua representations.

2 Three-Level Representation

Each lexical sense of a verb in our lexicon encodes its semantic meaning (i.e. semantic concept), its default predicate-argument mapping type (i.e. situation type), and any word-specific mapping exceptions (i.e. idiosyncrasies) in addition to its morphological and syntactic information. In the following, these three levels are discussed in detail.

2.1 Situation Types

Each of a verb's lexical senses is classified into one of the four default predicate-argument mapping types called *situation types*. As shown in Table 1, situation types of verbs are defined by two kinds of information: 1) the number of subcategorized NP or S arguments and 2) the types of thematic roles which these arguments should or should not map to. Since this kind of information is applicable to verbs of any language, situation types are language-independent

predicate-argument mapping types. Thus, in any language, a verb of type CAUSED-PROCESS has two arguments which map to AGENT and THEME in the default case (e.g. "kill"). A verb of type PROCESS-OR-STATE has one argument whose thematic role is THEME, and it does not allow AGENT as one of its thematic roles (e.g. "die"). An AGENTIVE-ACTION verb also has one argument but the argument maps to AGENT (e.g. "look"). Finally, an INVERSE-STATE verb has two arguments which map to THEME and GOAL; it does not allow AGENT for its thematic role (e.g. "see").

Although verbs in different languages are classified into the same four situation types using the same definition, mapping rules which map grammatical functions (i.e. subject, object, etc.) in the syntactic structures[2] to thematic roles in the semantic structures may be different from one language to another. This is because languages do not necessarily express the same thematic roles with the same grammatical functions. This mapping information is *language-specific* (cf. Nirenburg and Levin [21]).

The default mapping rules for the four situation types are shown in Table 2. They are nearly identical for the three languages (English, Spanish, and Japanese) we have analyzed so far. The only difference is that in Japanese the THEME of an INVERSE-STATE verb is expressed by marking the object NP with a particle "-ga", which is usually a subject marker (cf. Kuno [16]).[3] [4] So we add such information to the INVERSE-STATE mapping rule for Japanese. Generalization expressed in situation types has saved us from defining semantic mapping rules for each verb sense in each language, and also made it possible to acquire them from large corpora automatically.

This classification system has been partially derived from Vendler and Dowty's aspectual classifications [24, 12] and Talmy's lexicalization patterns [23]. For example, all AGENTIVE-ACTION verbs are so-called *activity* verbs, and so-called *stative* verbs fall under either INVERSE-STATE (if transitive) or PROCESS-OR-STATE (if intransitive).[5] However, the situation types are *not* for specifying the semantics of aspect, which is actually a property of the whole sentence rather than a verb itself (cf. Krifka [15], Dorr [11], Moens and Steedman [20]). For instance, as shown below, the same verb can be classified into two different aspectual classes (i.e. activity and accomplishment) depending on the types of object NP's or existence of certain PP's.

(1) a. Sue drank wine for/*in an hour.
    b. Sue drank a bottle of wine *for/in an hour.
(2) a. Harry climbed for/*in an hour.
    b. Harry climbed to the top *for/in an hour.

Situation types are intended to address the issue of cross-linguistic predicate-argument mapping generalization, rather than the semantics of aspect.

2.2 Semantic Concepts

Each lexical meaning of a verb is represented by a semantic concept (or frame) in our language-independent knowledge base, which is similar to the one described in Onyshkevych and Nirenburg [22]. Each verb frame has thematic role slots, which have two facets, TYPE and MAPPING. A TYPE facet value of a given slot provides a constraint on the type of objects which can be the value of the slot. This information is used during the semantic disambiguation process and during Debris Semantics, a semantically based syntax-to-semantics mapping strategy employed when a parse fails (cf. [1]).

In the MAPPING facets, we have encoded some cross-linguistically general predicate-argument mapping information. For example, we have defined that all the subclasses of #COMMUNICATION-EVENT# (e.g. #REPORT#, #CONFIRM#, etc.) map their sentential complements (SENT-COMP) to THEME, as shown below.

```
(#COMMUNICATION-EVENT#
      (AKO #DYNAMIC-SITUATION#)
      (AGENT (TYPE #PERSON# #ORGANIZATION#))
      (THEME (TYPE #SITUATION# #ENTITY#)
             (MAPPING (SENT-COMP T)))
      (GOAL  (TYPE #PERSON# #ORGANIZATION#)
             (MAPPING (P-ARG GOAL))))
```

Also, by grouping prepositions (for languages like English and Spanish) and postpositions (for those like Japanese) into semantically related concepts, we express other default mapping rules in the knowledge base language-independently. For example, the MAPPING facet of the INSTRUMENT slot in Figure 1 requires that the thematic role INSTRUMENT of the event #BREAK# be mapped from a P-ARG (i.e. a pre/postpositional argument in the syntactic structure) whose pre/postposition denotes INSTRUMENT. In the following three sentences in English,

[2] We use structures similar to LFG's f-structures. However, unlike those defined in Bresnan [4] or used in KBMT-89 [5], we normalize passive constructions.

[3] There is a debate over whether the NP with "ga" is a subject or object. However, our approach can accommodate either analysis.

[4] The GOAL of some INVERSE-STATE verbs in Japanese can be expressed by a "ni" postpositional phrase. However, as Kuno [16] points out, since this is an idiosyncratic phenomenon, such information does not go to the default mapping rule but to the lexicon.

[5] In both cases, the reverse does not hold.

	# of required NP or S arguments	default thematic roles	prohibited thematic roles
CAUSED-PROCESS	2	Agent Theme	-
PROCESS-OR-STATE	1	Theme	Agent
AGENTIVE-ACTION	1	Agent	-
INVERSE-STATE	2	Goal Theme	Agent

Table 1: Definitions of Situation Types

		English/Spanish Mapping	Japanese Mapping
CAUSED-PROCESS	AGENT THEME	(SURFACE SUBJECT) (SURFACE OBJECT)	(SURFACE SUBJECT) (SURFACE OBJECT)
PROCESS-OR-STATE	THEME	(SURFACE SUBJECT)	(SURFACE SUBJECT)
AGENTIVE-ACTION	AGENT	(SURFACE SUBJECT)	(SURFACE SUBJECT)
INVERSE-STATE	GOAL THEME	(SURFACE SUBJECT) (SURFACE OBJECT)	(SURFACE SUBJECT) (SURFACE OBJECT) (PARTICLE "GA")

Table 2: Default Mapping Rules for Three Languages

```
(#BREAK#
   (AKO #ANTI-CREATION-EVENT#)
   (AGENT (TYPE #CREATURE# #ORGANIZATION#))
   (THEME (TYPE #ENTITY#))
   (INSTRUMENT (TYPE #PHYSICAL-OBJECT#)
               (MAPPING (P-ARG INSTRUMENT))))
```

Figure 1: KB entry #BREAK#

Spanish and Japanese, the words which indicate the instrument used for the #BREAK# event are different (i.e. "with", "con", "-de"). However, by assigning the same semantic concept INSTRUMENT to each of these words in the lexicons, the same default mapping rule is shared across languages.

(3) John broke the vase *with* a hammer.
Juan rompió el vaso *con* un martillo.
John-wa kanazuchi-*de* kabin-o kowashita.

2.3 Idiosyncrasies

Idiosyncrasies slots in the lexicon specify word sense-specific idiosyncratic phenomena which cannot be captured by semantic concepts or situation types. In particular, subcategorized pre/postpositions of verbs are specified here. For example, the fact that "look" denotes its THEME argument by the preposition "at" is captured by specifying idiosyncrasies as shown in Figure 2. As discussed in the next section, we derive this kind of word-specific information automatically from corpora.

In addition, word-specific mapping rules are stated in idiosyncrasies slots to deal with thematic divergencies (e.g. (4)). In Figure 2, a Spanish entry "gustar" is shown.

(4) a. I like Mary.
b. María me gusta a mí.

Note that our representation can also handle cases where two internal arguments "flip-flop". For example, in (5), "infected" maps its object NP to GOAL and the PP to THEME idiosyncratically, while the Japanese counterpart "utsushita" maps its object NP ("infuruenza uirusu-o") to THEME and the PP ("saru-ni") to GOAL.[6] Such idiosyncratic mapping for "infect" is stated in the lexicon as in Figure 2.

(5) a. The researchers infected the monkeys with flu virus.
b. kenkyusha-tachi-wa saru-ni infuruenza uirusu-o utsushita.

Word-specific (hence language-specific) TYPE idiosyncrasies can be also stated in IDIOSYNCRASIES slots in the lexicon to deal with some of the semantic mismatches (cf. Barnett *et al.* [2], Kameyama *et al.* [14]). For example, Japanese verbs like "shibousuru (die)" and "shiyousuru (use)" take only people as their subjects while "shinu (die)" and "tsukau (use)" take any animate subjects. Such word-specific semantic constraint is defined in the lexicon as follows and overrides the constraint inherited from the knowledge base.

```
(SHIBOUSURU
   (CATEGORY . V)
   (SENSE-NAME . SHIBOUSURU-1)
   (INFLECTION-CLASS . IRS)
   (SEMANTIC-CONCEPT #DIE#)
   (IDIOSYNCRASIES (THEME (TYPE #PERSON#)))
   (SITUATION-TYPE PROCESS))
```

2.4 Sharing Information through Inheritance

The information stored in the three levels discussed above is shared through inheritance (cf. Jacobs [13]).

[6] The Japanese mapping is done by using the situation type of the verb (i.e. CAUSED-PROCESS) and the semantic concept assigned to the postposition "ni" (i.e. GOAL) (cf. Section 2.2).

```
(LOOK (CATEGORY . V)
      (SENSE-NAME . LOOK-1)
      (SEMANTIC-CONCEPT #LOOK#)
      (IDIOSYNCRASIES (THEME (MAPPING (P-ARG "AT"))))
      (SITUATION-TYPE AGENTIVE-ACTION))

(GUSTAR (CATEGORY . V)
        (SENSE-NAME . GUSTAR-1)
        (SEMANTIC-CONCEPT #LIKE#)
        (IDIOSYNCRASIES (THEME (MAPPING (SURFACE SUBJECT)))
                        (GOAL (MAPPING (SURFACE OBJECT))))
        (SITUATION-TYPE INVERSE-STATE))

(INFECT (CATEGORY . V)
        (SENSE-NAME . INFECT-1)
        (SEMANTIC-CONCEPT #INFECT#)
        (IDIOSYNCRASIES (THEME (MAPPING (P-ARG "WITH")))
                        (GOAL (MAPPING (SURFACE OBJECT))))
        (SITUATION-TYPE CAUSED-PROCESS))
```

Figure 2: Lexical entries for "look", "gustar", and "infect"

```
(BREAK.1 (ISA #BREAK#)
         (THEME WINDOW.2)
         (TENSE PAST))
```

Figure 4: "The window broke."

```
(BREAK.5 (ISA #BREAK#)
         (AGENT JOHN.6)
         (THEME WINDOW.7)
         (TENSE PAST))
```

Figure 5: "John broke the window."

The idiosyncrasies take precedence over the situation types and the semantic concepts, and the situation types over the semantic concepts. For example, two mapping rules derived by inheritance for "break" are shown in Figure 3. Notice that the semantic TYPE restriction and INSTRUMENT role stored in the knowledge base are also inherited at this time.

By separating language-specific predicate-argument mapping information (i.e. situation types and idiosyncrasies) from semantic concepts, we can capture predicate-argument mapping divergences without proliferating semantic concepts. For instance, we generate Japanese sentences from English within this framework in the following way. Given a semantic representation for the sentence "The window broke" (cf. Figure 4, detail omitted), a verb is chosen from the Japanese lexicon whose semantic concept is #BREAK# and which expresses its THEME but not its AGENT, namely "kowareru", which has situation type PROCESS-OR-STATE. On the other hand, given a semantic representation for the sentence "John broke the window" (cf. Figure 5), a verb whose semantic concept maps to #BREAK# and which expresses both its AGENT and THEME, namely "kowasu", which has situation-type CAUSED-PROCESS, is chosen.

3 Automatic Acquisition from Corpora

In order to expand our lexicon to the size needed for broad coverage and to be able to tune the system to specific domains quickly, we have implemented algorithms to automatically build multi-lingual lexicons from corpora. The three-level representation described here lends itself to automatic acquisition. In this section, we discuss how the situation types and lexical idiosyncrasies are determined for verbs.

Our overall approach is to use simple robust parsing techniques that depend on a few language-dependent syntactic heuristics (e.g. in English and Spanish, a verb's object usually directly follows the verb), and a dictionary for part of speech information. We have used these techniques to acquire information from English, Spanish, and Japanese corpora varying in length from about 25000 words to 2.7 million words.

3.1 Acquiring Situation Type Information

We use two surface features to restrict the possible situation types of a verb: the verb's *transitivity rating* and its *subject animacy*.

The transitivity rating of a verb is defined to be

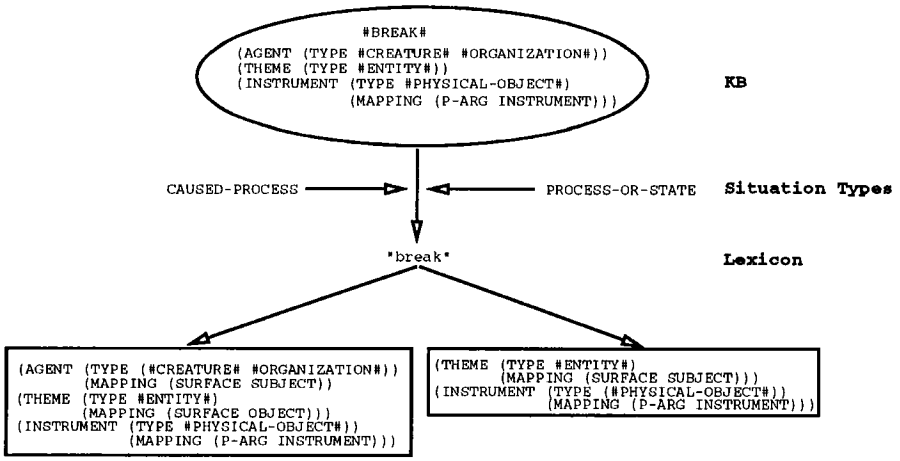

Figure 3: Information from the KB, the situation type, and the lexicon all combine to form two predicate-argument mappings for the verb "break."

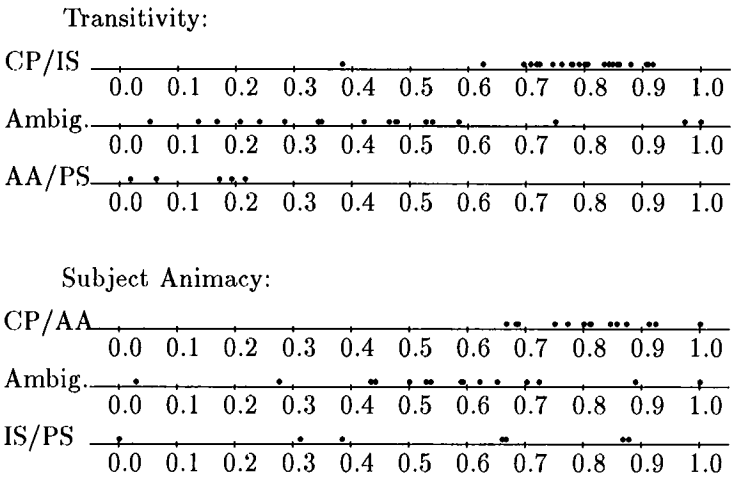

Figure 6: This graph shows the accuracy of the Transitivity and Subject Animacy metrics.

verb	occs	TR	SA	Pred. ST	Correct ST	Prepositional Idio
SUFFICE	8	0.6250	0.0000	(IS)	(IS)	
TIME	15	0.8333	1.0000	(CP IS)	(CP)	
TRAIN	20	1.0000	1.0000	(CP IS)	(CP PS)	at
WRAP	22	0.7222	0.6667	(CP IS)	(CP)	up over in with
SORT	25	0.4211	1.0000	(CP IS AA PS)	(CP AA)	out
UNITE	27	0.5833	1.0000	(CP IS AA PS)	(CP AA)	
TRANSPORT	28	0.8571	0.6667	(CP IS)	(CP)	
SUSTAIN	32	0.9062	0.6842	(CP IS)	(CP)	
SUBSTITUTE	33	0.7500	0.5000	(IS)	(CP PS)	for
TARGET	36	0.7778	0.8000	(CP IS)	(CP)	
STORE	36	0.9091	1.0000	(CP IS)	(CP)	on
STEAL	36	0.9167	0.6667	(CP IS)	(CP)	from
SHUT	36	0.2400	0.5000	(IS PS)	(CP PS)	up for
STRETCH	53	0.5278	0.5000	(IS PS)	(CP PS)	over into out from
STRIP	57	0.7609	0.8571	(CP IS)	(CP)	from into of
THREATEN	58	0.8793	0.4419	(IS)	(CP IS)	over
WEAR	61	0.8033	0.6667	(CP IS)	(IS)	over
TREAT	77	0.8052	0.8000	(CP IS)	(CP)	as
TERMINATE	79	0.9726	1.0000	(CP IS)	(CP PS)	
WEIGH	81	0.2069	0.5294	(IS PS)	(CP PS)	on with into
TEACH	82	0.7794	0.6875	(CP IS)	(CP)	at
SURROUND	85	0.8000	0.6667	(CP IS)	(CP)	
TOTAL	97	0.0515	0.2759	(PS)	(CP PS)	at
VARY	112	0.1354	0.0294	(IS PS)	(CP PS)	from over
WAIT	130	0.1923	1.0000	(CP IS AA PS)	(AA)	for up
SPEAK	139	0.1667	0.7500	(CP IS AA PS)	(AA CP)	out at up
SURVIVE	146	0.4754	0.3846	(IS PS)	(IS PS)	
UNDERSTAND	180	0.6946	0.8684	(CP IS)	(IS)	
SURGE	188	0.0182	0.3125	(PS)	(PS)	
SUPPLY	188	0.7176	0.8571	(CP IS)	(CP)	with
SIT	199	0.0625	0.7027	(AA PS)	(AA PS)	on with at out in up
TEND	200	0.8594	0.4340	(IS)	(CP IS)	
BREAK	219	0.4771	0.5000	(IS PS)	(CP PS)	up into out
WRITE	243	0.4637	0.9123	(CP IS AA PS)	(CP AA)	off
WATCH	268	0.7069	0.8462	(CP IS)	(CP)	out over
SUCCEED	277	0.5379	0.8899	(CP IS AA PS)	(CP PS)	
STAY	300	0.2156	0.6604	(CP IS AA PS)	(PS)	out up on with at
STAND	310	0.2841	0.7237	(CP IS AA PS)	(PS CP AA)	up at as out on
TELL	368	0.8054	0.8101	(CP IS)	(CP)	
SPEND	445	0.3823	0.8125	(CP IS AA PS)	(CP)	on over
SUPPORT	454	0.8486	0.5370	(IS)	(CP IS)	
SUGGEST	570	0.7782	0.5918	(IS)	(CP IS)	
TURN	852	0.3418	0.5891	(IS PS)	(CP PS)	out into up over
START	890	0.3474	0.6221	(CP IS AA PS)	(CP PS)	with off out
LOOK	1084	0.1718	0.6520	(CP IS AA PS)	(AA PS)	at into for up
THINK	1227	0.7602	0.9237	(CP IS)	(CP)	
TRY	1272	0.7904	0.8743	(CP IS)	(CP)	
WANT	1659	0.8559	0.8787	(CP IS)	(IS)	
USE	2211	0.8416	0.7725	(CP IS)	(CP)	
TAKE	2525	0.7447	0.5933	(IS)	(CP IS)	over off out into up

Table 3: Automatically Derived Situation Type and Idiosyncrasy Data

the number of transitive occurrences in the corpus divided by the total occurrences of the verb. In English, a verb appears in the transitive when either:

- the verb is directly followed by a noun, determiner, personal pronoun, adjective, or wh-pronoun (e.g. "John owns a cow.")
- the verb is directly followed by a "THAT" as a subordinating conjunction (e.g. "John said that he liked cheese.")
- the verb is directly followed by an infinitive (e.g. "John promised to walk the dog.")
- the verb past participle is preceded by "BE," as would occur in a passive construction (e.g. "The apple was eaten by the pig.")

For Spanish, we use a very similar algorithm, and for Japanese, we look for noun phrases with an object marker near and to the left of the verb. A high transitivity is correlated with CAUSED-PROCESS and INVERSE-STATE while a low transitivity correlates with AGENTIVE-ACTION and PROCESS-OR-STATE. Table 3 shows 50 verbs and their calculated transitivity rating. Figure 6 shows that for all but one of the verbs that are unambiguously transitive the transitivity rating is above 0.6. The verb "spend" has a transitivity rating of 0.38 because most of its direct objects are numeric dollar amounts. Phrases which begin with a number are not recognized as direct objects, since most numeric amounts following verbs are adjuncts as in "John ran 3 miles."

We define the a verb's subject animacy to be the number of times the verb appears with an animate subject over the total occurrences of the verb where we identified the subject. Any noun or pronoun directly preceding a verb is considered to be its subject. This heuristic fails in cases where the subject NP is modified by a PP or relative clause as in "The man under the car wore a red shirt." We have only implemented this metric for English. The verb's subject is considered to be animate if it is any one of the following:

- A personal pronoun ("it" and "they" were excluded, since they may refer back to inanimate objects.)
- A proper name
- A word under "agent" or "people" in WordNet (cf. [19])
- A word that appears in a MUC-4 template slot that can be filled only with humans (cf. [9])

Verbs that have a low subject animacy cannot be either CAUSED-PROCESS or AGENTIVE-ACTION, since the syntactic subject must map to the AGENT thematic role. A high subject animacy does not correlate with any particular situation type, since several stative verbs take only animate subjects (e.g. perception verbs).

The predicted situation types shown in Figure 6 were calculated with the following algorithm:

1. Assume that the verb can occur with every situation type.
2. If the transitivity rating is greater than 0.6, then discard the AGENTIVE-ACTION and PROCESS-OR-STATE possibilities.
3. If the transitivity rating is below 0.1, then discard the CAUSED-PROCESS and INVERSE-STATE possibilities.
4. If the subject animacy is below 0.6, then discard the CAUSED-PROCESS and AGENTIVE-ACTION possibilities.

We are planning several improvements to our situation type determination algorithms. First, Because some stative verbs can take animate subjects (e.g. perception verbs like "see", "know", etc.), we sometimes cannot distinguish between INVERSE-STATE or PROCESS-OR-STATE and CAUSED-PROCESS or AGENTIVE-ACTION verbs. This problem, however, can be solved by using algorithms by Brent [3] or Dorr [11] for identifying stative verbs.

Second, verbs ambiguous between CAUSED-PROCESS and PROCESS-OR-STATE (e.g. "break", "vary") often get inconclusive results because they appear transitively about 50% of the time. When these verbs are transitive, the subjects are almost always animate and when they are intransitive, the subjects are nearly always inanimate. We plan to recognize these situations by calculating animacy separately for transitive and intransitive cases.

3.2 Acquiring Idiosyncratic Information

We automatically identify likely pre/postpositional argument structures for a given verb by looking for pre/postpositions in places where they are likely to attach to the verb (i.e. within a few words to the right for Spanish and English, and to the left for Japanese). When a particular pre/postposition appears here much more often than chance (based on either Mutual Information or a chi-squared test [7, 6]),

word	possible clausal complements
know	THATCOMP
vow	THATCOMP, TOCOMP
eat	–
want	TOCOMP
resume	INGCOMP

Table 4: English Verbs which Take Complementizers

we assume that it is a likely argument. A very similar strategy works well at identifying verbs that take sentential complements by looking for complementizers (e.g. "that", "to") in positions of likely attachment. Some examples are shown in Tables 3 and 4. The details of the exact algorithm used for English are contained in McKee and Maloney [18]. Areas for improvement include distinguishing between cases where a verb takes a prepositional arguments, a prepositional particle, or a common adjunct.

4 Conclusion

We have used this three-level knowledge representation for our English, Spanish and Japanese lexicons, and have built lexicons with data derived from corpora for each language. The lexicons have been used for several multilingual data extraction applications (cf. Aone *et al.* [1]) and a prototype MT system (Japanese - English). The fact that this representation is language-independent, bidirectional and amenable to automatic acquisition has minimized our data acquisition effort considerably.

Currently we are investigating ways in which thematic role slots of verb frames and semantic type restrictions on these slots are derived automatically from corpora (cf. Dagan and Itai [8], Zernik and Jacobs [25]) so that knowledge acquisition at all three levels can be automated.

References

[1] Chinatsu Aone, Doug McKee, Sandy Shinn, and Hatte Blejer. SRA: Description of the SOLOMON System as Used for MUC-4. In *Proceedings of Fourth Message Understanding Conference (MUC-4)*, 1992.

[2] Jim Barnett, Inderjeet Mani, Elaine Rich, Chinatsu Aone, Kevin Knight, and Juan C. Martinez. Capturing Language-Specific Semantic Distinctions in Interlingua-based MT. In *Proceedings of MT Summit III*, 1991.

[3] Michael Brent. Automatic Semantic Classification of Verbs from Their Syntactic Contexts: An Implemented Classifier for Stativity. In *Proceedings of the 5th European ACL Conference*, 1991.

[4] Joan Bresnan, editor. *The Mental Representation of Grammatical Relations.* MIT Press, 1982.

[5] Center for Machine Translation, Carnegie Mellon University. *KBMT-89: Project Reprot*, 1989.

[6] Kenneth Church and William Gale. Concordances for Parallel Text. In *Proceedings of the Seventh Annual Conference of the University of Waterloo Centre for the New OED and Text Research: Using Corpora*, 1991.

[7] Kenneth Church and Patrick Hanks. Word Association Norms, Mutual Information, and Lexicography. *Computational Linguistics*, 16(1), 1990.

[8] Ido Dagan and Alon Itai. Automatic Acquisition of Constraints for the Resolution of Anaphora References and Syntactic Ambiguities. In *Proceedings of the 13th International Conference on Computational Linguistics*, 1990.

[9] Defense Advanced Research Projects Agency. *Proceedings of Fourth Message Understanding Conference (MUC-4).* Morgan Kaufmann Publishers, 1992.

[10] Bonnie Dorr. Solving Thematic Divergences in Machine Translation. In *Proceedings of 28th Annual Meeting of the ACL*, 1990.

[11] Bonnie Dorr. A Parameterized Approach to Integrating Aspect with Lexical-Semantics for Machine Translation. In *Proceedings of 30th Annual Meeting of the ACL*, 1992.

[12] David Dowty. *Word Meaning and Montague Grammar.* D. Reidel, 1979.

[13] Paul Jacobs. Integrating Language and Meaning in Structured Inheritance Networks. In John Sowa, editor, *Principles of Semantic Networks.* Morgan Kaufmann, 1991.

[14] Megumi Kameyama, Ryo Ochitani, and Stanley Peters. Resolving Translation Mismatches with Information Flow. In *Proceedings of 29th Annual Meeting of the ACL*, 1991.

[15] Manfred Krifka. Nominal Reference, Temporal Construction, and Quantification in Event Semantics. In R. Bartsch et al., editors, *Semantics*

and Contextual Expressions. Foris, Dordrecht, 1989.

[16] Susumu Kuno. *The Structure of the Japanese Language.* MIT Press, 1973.

[17] Michael McCord. Design of LMT: A prolog-based Machine Translation System. *Computational Linguistics*, 15(1), 1989.

[18] Doug McKee and John Maloney. Using Statistics Gained from Corpora in a Knowledge-Based NLP System. In *Proceedings of The AAAI Workshop on Statistically-Based NLP Techniques*, 1992.

[19] George Miller, Richard Beckwith, Christiane Fellbaum, Derek Gross, and Katherine Miller. Five papers on WordNet. Technical Report CSL Report 43, Cognitive Science Laboratory, Princeton University, 1990.

[20] Marc Moens and Mark Steedman. Temporal ontology and temporal reference. *Computational Linguistics*, 14(2), 1988.

[21] Sergei Nirenburg and Lori Levin. Syntax-Driven and Ontology-Driven Lexical Semantics. In *Proceedings of ACL Lexical Semantics and Knowledge Representation Workshop*, 1991.

[22] Boyan Onyshkevych and Sergei Nirenburg. Lexicon, Ontology and Text Meaning. In *Proceedings of ACL Lexical Semantics and Knowledge Representation Workshop*, 1991.

[23] Leonard Talmy. Lexicalization Patterns: Semantic Structure in Lexical Forms. In Timothy Shopen, editor, *Language Typology and Syntactic Descriptions.* Cambridge University Press, 1985.

[24] Zeno Vendler. *Linguistics in Philosophy.* Cornell University Press, 1967.

[25] Uri Zernik and Paul Jacobs. Tagging for Learning: Collecting Thematic Relations from Corpus. In *Proceedings of the 13th International Conference on Computational Linguistics*, 1990.