Semantic Analysis in The Mikrokosmos Machine Translation Project

Stephen Beale, Sergei Nirenburg and Kavi Mahesh

Computing Research Laboratory
Box 30001, Dept. 3CRL
New Mexico State University
Las Cruces, NM 88003-0001 USA
sb,sergei,mahesh@crl.nmsu.edu

May 17, 1996

Abstract

The Mikrokosmos (uK) Machine Translation System is a knowledge-based machine translation (KBMT) system under development at New Mexico State University. Unlike previous research in interlingual MT, this project is a large-scale, practical MT system. By the end of the year, a lexicon of approximately 20,000 Spanish words (with over 35,000 word senses) supported by an ontology of 6000 concepts will be in place. High quality semantic analyses of over 400 article-length Spanish texts in the domain of company mergers and acquisitions will have been produced. In the coming year, uK intends to expand into other languages, notably Thai.

This paper introduces the central concepts involved in KBMT, including Text-Meaning Representation (TMR), ontology, and the semantic lexicon. The semantic analyzer ''engine'' will then be described in detail, with examples of how knowledge from the ontology, lexicon and syntactic analysis are combined to create the basic semantic dependency structures found in the TMR outputs. Several issues in semantic analysis programming will be briefly probed, including dependency-directed processing, ''best-first'' search, and knowledgeable treatment of unexpected and ambiguous inputs.
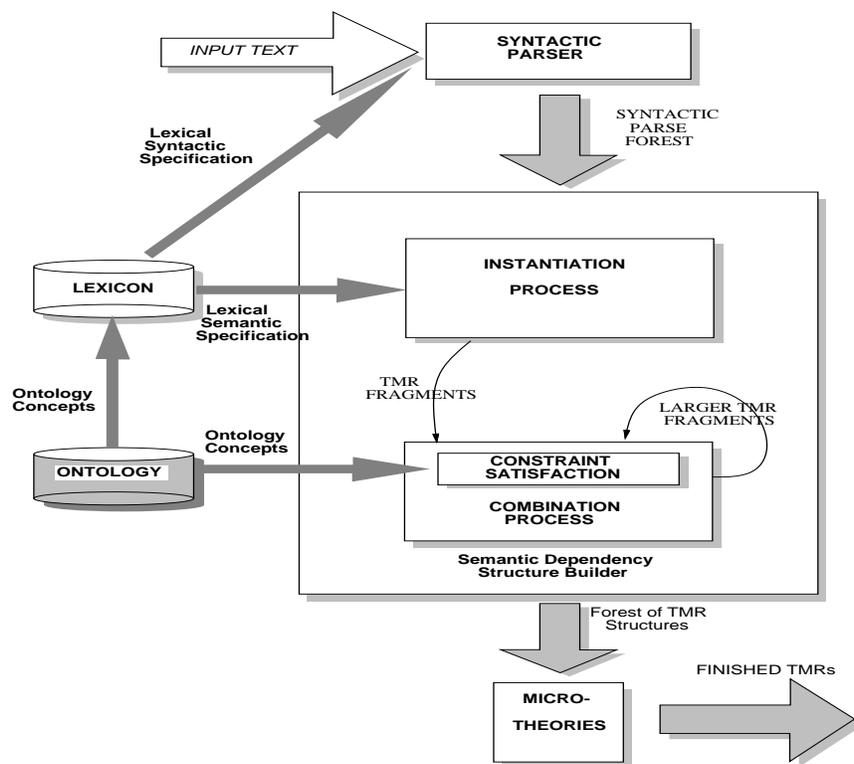
Figure 1: The Mikrokosmos NLP Architecture.

1 The Mikrokosmos Machine Translation Project

A comprehensive study of the computational treatment of texts is a multifaceted endeavor covering a wide range of linguistic and language use phenomena. Because the various facets of this knowledge are complex in their own right, study of any individual phenomenon is often conducted in relative isolation from the study of other related phenomena. However, in a knowledge-based machine translation (KBMT) application, knowledge about a large number of interrelated linguistic and language use phenomena is required. A natural way of combining the diverse knowledge required of such a system into a unified whole is for the various phenomena to be treated by separate computational linguistic "microtheories" united through a system's control architecture and knowledge representation conventions.

In the uK project, being developed by researchers at the Computing Research Laboratory (CRL) of New Mexico State University,[1] a comprehensive study of a variety of microtheories central to the support of KBMT systems is being carried out with the ultimate objective of defining a methodology for representing the meaning of natural language texts in a language-neutral interlingual format called a text meaning representation (TMR). The TMR represents the result of analysis of a given input text in any one of the languages supported by the KBMT system, and serves as input to the generation process. The meaning of the input text is represented in the TMR as elements of an independently motivated model of the world (or ontology). The link between the ontology and the TMR is provided by the lexicon, where the meanings of most open class lexical items are defined in terms of their mappings into ontological concepts and their resulting contributions to TMR structure. Information about the nonpropositional components of text meaning such as speech acts, speaker attitudes and intentions, relations among text units, deictic references, etc. is also derived from the lexicon with inputs from other microtheories, and becomes part of the TMR. Figure 1 illustrates the uK architecture for analyzing input texts.

Initially, the project is concentrating on the microtheory of lexical-semantic dependency, the core microtheory underlying our approach to a comprehensive analysis of the meaning of texts, and the one in which the basic structure of events or states and their properties is specified. Additional microtheories are being developed for aspect, time,

---

[1] Please see the CRL WWW home page for a more complete overview of the uK Project at http://crl.nmsu.edu/index.html
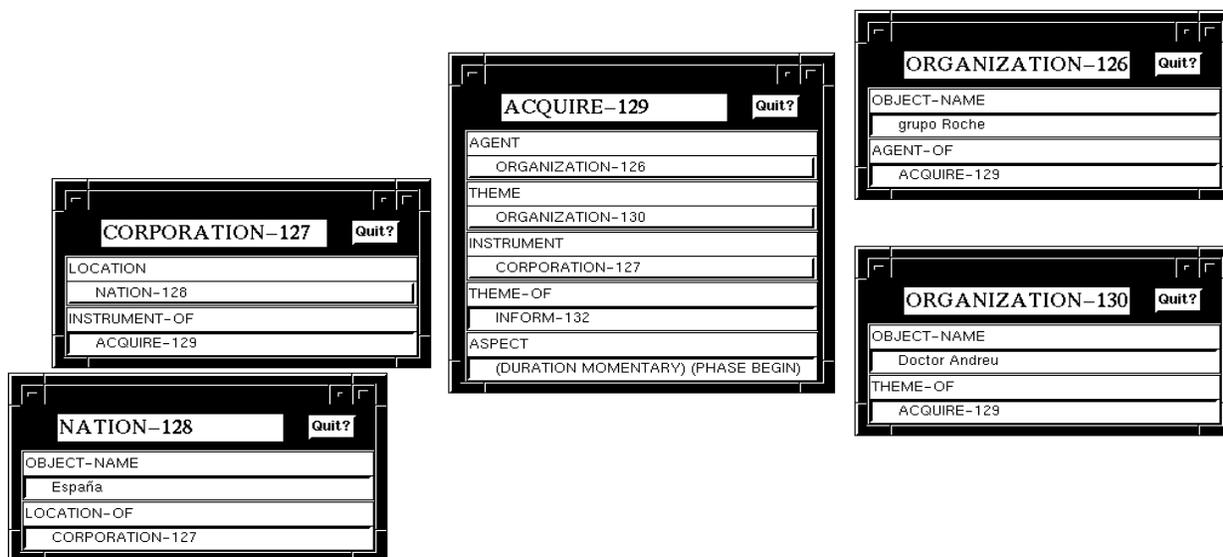
Figure 2: Partial Text Meaning Representation of Example Sentence.

modalities, discourse relations, reference, event ellipsis and style.[2]

2   Text Meaning Representations

A TMR is a language-neutral description (an interlingua) of the meaning conveyed in a natural language text, and is derived by syntactic, semantic, and pragmatic analysis of the text. Because the TMR is intended to be language neutral, it is also deliberately syntax neutral, and avoids using terminology like clause, proposition, tense, etc., which are associated more closely with the syntactic structure of a particular language. In addition to providing information about the lexical-semantic dependencies in the text, the TMR represents stylistic factors, discourse relations, speaker attitudes, and other pragmatic factors present in the discourse structure. In doing so, the TMR captures not only the meaning of individual elements in the text, but also the relations between those elements, and captures both propositional and nonpropositional components of textual meaning.

The results of analysis of an input text are represented in a formal, frame-based language. The meanings of most open-class lexical units are represented by instantiating, combining and constraining concepts available in the ontology. However, the intent of a text cannot fully be captured by instantiating ontological concepts alone; information about pragmatic and discourse related phenomena must be represented, and relations between components of meaning must also be expressed. To facilitate this, the TMR language contains special notation for representing attitudes, relations, speech acts, time, quantities, rates, and sets.

Figure 2 displays a portion of the TMR output for sentence 1:

1a. El grupo Roche, a traves de su compania en Espana, adquirio Doctor Andreu, se informo hoy aqui.
1b. The Roche group, through its company in Spain, acquired Doctor Andrew, it was announced today.

The central concept for the ''acquire'' clause is ACQUIRE-129. This maps various concepts into the AGENT, THEME and INSTRUMENT slots. The significance of these mappings and how they were selected will be detailed below.

[2]For example, see (Viegas and Nirenburg 1995) for a treatment of verbal ellipsis.
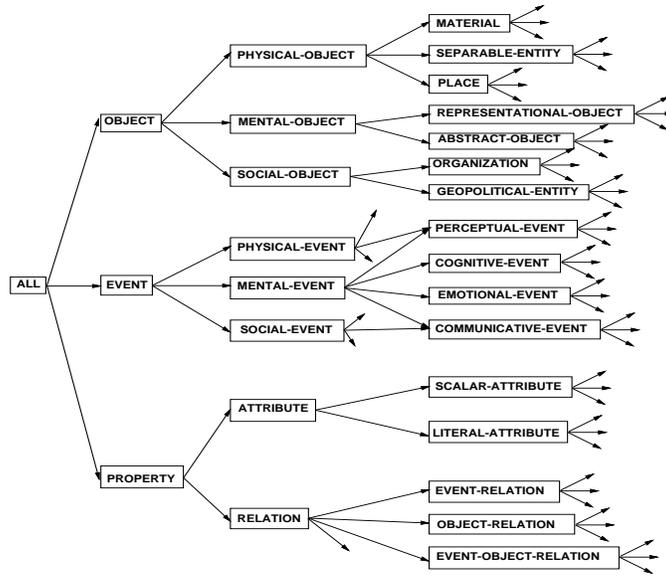
Figure 3: Top-Level Hierarchy of the Mikrokosmos Ontology Showing the First Three Levels of the Object, Event, and Property Taxonomies.

3 Ontology

An ontology for NLP purposes is a body of knowledge about the world (or a domain) that a) is a repository of primitive symbols used in meaning representation; b) organizes these symbols in a tangled subsumption hierarchy; and c) further interconnects these symbols using a rich system of semantic relations defined among the concepts. In order for such an ontology to become a computational resource for solving problems such as ambiguity and reference resolution, it must be actually constructed, not merely defined formally. The ontology must be put into well-defined relations with other knowledge sources in the system. In this application, the ontology supplies world knowledge to lexical, syntactic and semantic processes, and other microtheories.

We are currently in the process of a massive acquisition of objects, events and properties related to the domain of company mergers and acquisitions (Mahesh and Nirenburg, 1995). Over the period of about three months, the uK ontology has acquired over 2000 concepts organized in a tangled hierarchy with ample interconnection across the branches. The ontology emphasizes depth in organizing concepts and reaches depth 10 or more along a number of paths. The branching factor is kept much less than 10 at most points. Each concept has, on average, 5 to 10 slots linking it to other concepts or literal constants. The top levels of the hierarchy (Figure 3) have proved very stable as we are continuing to acquire new concepts at the lower levels.

Unlike many other ontologies with a narrow focus, our ontology must cover a wide variety of concepts in the world. In particular, our ontology cannot stop at organizing terminological nouns into a taxonomy of objects and their properties; it must also represent a taxonomy of (possibly, complex) events and include many interconnections between objects and events to support a variety of disambiguation tasks. For example, in the sample text above, the analyzer must distinguish between two meanings of ''adquirir'': 1) ACQUIRE, and 2) LEARN, where ACQUIRE and LEARN are concepts in the ontology defined in Figure 4.

In our example sentence, the fact that the THEME of LEARN is constrained to be INFORMATION will be enough to eliminate it from consideration. Additional examples of disambiguation will be given below.

The ontology aids natural language processing in the following ways:

- It represents selectional preferences for relations between concepts. This knowledge is invaluable for resolving ambiguities by means of the constraint satisfaction process.

- It enables inferences to be made from the input text using knowledge contained in the concepts. This can help resolve ambiguities as well as fill gaps in the text meaning. A default value from the ontological concept can be

Figure 4: Ontological Definition of ACQUIRE and LEARN.

filled in a slot, for example, when a text does not provide a specific value.

- It enables inferences to be made using the topology of the network, as in searching for the shortest path between two concepts. Such search-based inferences can support metonymy and metaphor processing, figuring out the meaning of a complex nominal or be used in constraint relaxation when the input cannot be treated with the available knowledge.

4 Semantic Lexicon

In the model of NLP adopted in a KBMT paradigm, the lexicon becomes the key locus and source of knowledge. Compared to many other computational lexicons, in our approach a substantial amount of information is either directly located in the lexicon, or is indexed or referenced through the lexicon. Figure 5 depicts the lexicon entries for the spanish word ''adquirir''. There are two entries corresponding to the two word-senses we currently identify. Sense 1 maps into an ACQUIRE concept while sense 2 maps into a LEARN concept.

An entry in the lexicon is comprised of a number of zones, integrating various levels of lexical information, from phonological and morphological to lexical-semantic and pragmatic information. Of particular interest to us here are the SYN-STRUC and SEM zones.

4.1 SYN-STRUC Zone

The content of the SYN-STRUC zone of a lexicon entry is an indication of where the lexeme may fit into the syntactic parse of a sentence. In addition, this zone provides the basis of the syntax-semantics interface. The information contained in this zone essentially amounts to an underspecified piece of a syntactic parse of a typical sentence using the lexeme; this piece contains the lexeme in question, and may include information from any number of embedded levels (but typically not more than two) above or below the current lexeme. The information included reflects those

5
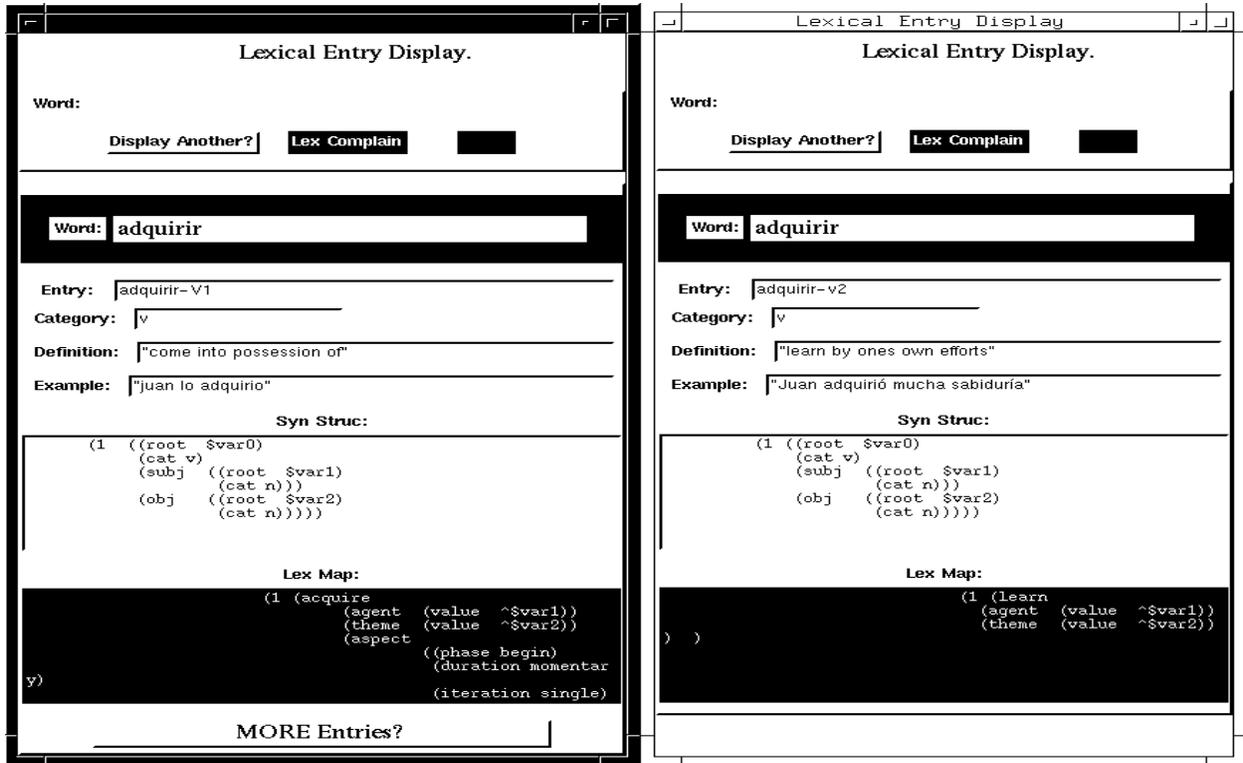
Figure 5: Lexicon Entries for ''adquirir''.

levels and elements of the syntactic parse which the current lexeme syntactically selects for; in the current model, verbs select for all their arguments, modifiers select for their heads, prepositions select for their objects as well as for their node of attachment, etc. The SYN-STRUC zone thus determine such things as subcategorization (optionality is indicated, otherwise obligatoriness assumed), complements allowed, etc.

In the SYN-STRUC zone, we place variables at the ROOT positions selected for by the lexeme in question, which is identified by the variable var0. Subsequently numbered variables (var1, var2, ...) identify other syntactic nodes with which the current lexeme has syntactic or semantic dependencies. For example, the pattern below is appropriate for any regular monotransitive verb:[3]

```
((root var0)
 (subj ((root var1) (cat n)))
 (obj ((root var2) (cat n))))
```

For instance, in our example sentence, while processing the ''adquirir'' lexicon entry, ''Grupo Roche'' will be bound to var1 as the SUBJ, while ''Dr. Andreu'' will be bound to var2. If a SYN-STRUC requires a syntactic pattern not found in the current sentence, then that word sense is not used.[4]

The variable bindings introduced in the SYN-STRUC provide an interaction with the meaning pattern from the SEM zone in that certain portions of the meaning pattern for a phrase or clause are regularly and compositionally determined by the semantics of the components (Principle of Compositionality); the structure of the resulting meaning pattern is determined not only by the semantic meaning patterns of each of the components, but also by their syntactic relationship in the SYN-STRUC zone.

---

[3] An LFG-like syntactic description is used.

[4] Except possibly in failure recovery, described below.

Grupo-Roche	a-traves-de	su	compania	en	espana	adquirir	Dr. Andrew
ORGANIZATION	*LOCATION*	*OWNER*	*CORPORATION*	*LOCATION*	*NATION*	*ACQUIRE*	*HUMAN*
	INSTRUMENT		*SOCIAL-EVENT*	*TEMPORAL*		*LEARN*	*ORGANIZATION*

Figure 6: Possible Word Senses for Example Sentence.

4.2 SEM Zone

The SEM zone provides the mapping to the output semantics. Each SEM zone is basically an underspecified TMR fragment which includes as much meaning as can be extracted from the word being processed. The interaction of SEM zones from all the words in the sentence[5] result in the final TMR outputs.

Referring to Figure 5 again, the adquirir-v1 SEM zone creates an ACQUIRE concept with AGENT and THEME slots that will be filled by the TMR names that are produced by ''grupo Roche'' (var1) and ''Dr. Andreu'' (var2), respectively. Other words in the sentence can fill in additional information in the ACQUIRE TMR. One of the meanings of ''a traves de,'' treated as a phrasal entry, will add an INSTRUMENT slot.

In addition to specifying TMR fragments, the SEM zone can add in language-specific semantic constraints which add to or override the language-neutral constraints provided by the ontology.[6] For example, the English verb ''to taxi,'' as in ''the jet taxied to the gate'' maps into a GROUND-CONTACT-MOTION, but further specifies that its INSTRUMENT must have AIRCRAFT semantics. These ''constrained mappings'' from language-specific definitions to language-neutral concepts arise because the ontology does not attempt to provide concepts for every conceivable event,[7] nor is its goal to predict all of the idiosyncratic constraints found in different natural languages.[8]

5 The Semantic Analyzer

The semantic analyzer is charged with the task of combining the knowledge contained in the ontology and lexicon and applying it to the current input to produce output TMRs. The central tasks involved in this endeavor are to retrieve the appropriate semantic constraints for each possible word sense, test each in context, and construct the output TMRs by instantiating the SEM zones of the word senses which, taken together, best satisfy the combination of constraints. Below, we will examine the steps taken to choose the ACQUIRE meaning of ''adquirir'' over the LEARN meaning. We will then briefly trace out the other decisions made and provide a summary of the computational methods applied in the analysis process.

5.1 Generating Constraints

Step 0 in the semantic analysis process is acquiring the syntactic analysis of the input sentence. To avoid duplication of effort, uK uses the output of the Pangloss MT syntactic analysis module (or Panglyzer), also developed at NMSU.[9] Since, at the present time, the Panglzer makes all attachment decisions, uK is limited to deciding between word sense meanings.[10]

The first real step for uK is to gather up all of the possible lexicon entries for each of the words. Figure 6 lists all the word-sense mapping possibilities for the example sentence. For ''adquirir,'' the two lexicon entries shown in Figure 5 are retrieved, with mappings into ACQUIRE and LEARN word senses. For each word sense, the SYN-STRUC zone must be examined to see if it fits the current sentence. If it does, then the VARs must be bound to their corresponding

---

[5]along with information added by other microtheories

[6]The next section describes how the analyzer retrieves and applies constraints from the ontology

[7]For example, a South-American Indian language has a single word for ''she carries water down to the river'', but our ontology sadly cannot map directly into such an event.

[8]For example, one word for a human drinking, another word for animals drinking.

[9]We would prefer to interleave semantics in the syntactic analysis process. Currently, we are investigating ways to modify the Pangloss-uK interface to provide a level of interleaving, especially with regards to PP attachments.

[10]Again, this limitation will be removed shortly. Choosing between attachments will proceed in the same constraint-satisfaction paradigm as described for word senses, with some possible inputs from attachment microtheories such as ''minimal attachment'', etc..

```
      "adquirir" SYN-STRUC                              Input Syntax

((root $VAR0)         ←─────────────────────→  ((root ADQUIRIR-1)
 (cat v)              ←·········································→   (cat v)
 (subj ((root $VAR1)  ←─────────────────────→   (subj ((root GRUPO-ROCHE-2)
        (cat n)))     ←·········································→          (cat n)))
 (obj  ((root $VAR2)  ←─────────────────────→   (obj  ((root DR-ANDREW-3)
        (cat n)))     ←·········································→          (cat n)))
 )                                              (pp-adjunct ((root A-TRAVES-DE-4)
                                                     (obj  ((root COMPANIA-5)
                                                        (mod ((root SU-6)))
                                                        (pp-adjunct ((root EN-7)
                                                            (obj ((root ESPANA-8)))))))))
                                                 )
```
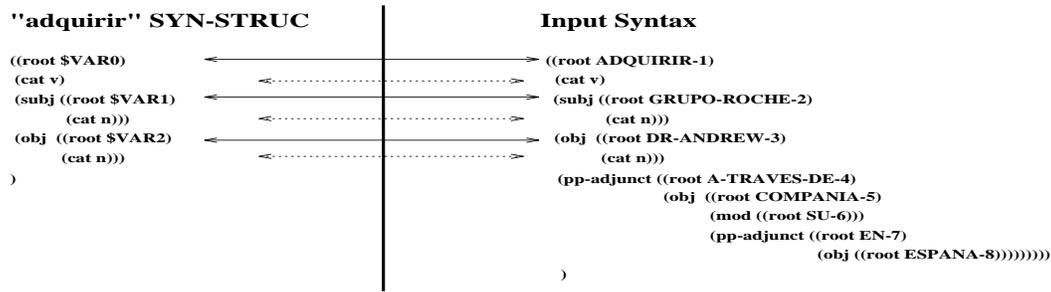
Figure 7: Variable binding for ''adquirir''.

word instances in the current input sentence. For ''adquirir,'' both word senses have identical SYN-STRUC zones, so the variable binding process displayed in Figure 7 applies to both.

After variable binding, the semantic analyzer examines the SEM zone of each word sense in order to construct a list of constraints that must be satisfied for that word sense. Constraints can arise from five sources:

1. The ontological definition of the current word-sense restricts the semantics of its slot fillers. The definitions for ACQUIRE and LEARN are shown in Figure 4. ACQUIRE and LEARN both require a HUMAN AGENT. ACQUIRE requires a non-HUMAN OBJECT for its THEME, while LEARN requires an INFORMATION THEME.

2. The ontological definition of the word-sense that will **fill** the slot restricts the kind of slots it may be the filler of. Type 1 constraints ask ''What kind of fillers do I allow?'' Type 2 constraints ask about the **fillers**, ''What kind of concepts can this filler modify with the given slot?'' For instance, HAMMER, when used as the filler for an INSTRUMENT slot usually modifies some sort of BUILD event. In the example, ORGANIZATION (from Grupo-Roche-1) as an AGENT filler currently[11] does not select for any specific type of event, nor do ORGANIZATION (Dr-Andrew-1) or HUMAN (Dr-Andrew-2) as THEMEs select for a specific event.

3. The ontological definition of the **slot** (the property name that is being added) restricts what its DOMAIN and RANGE can be. Sometimes, in the absence of more specific constraints from 1 and 2 above, uK can find default values by looking up the slot itself in the ontology. An AGENT slot requires its DOMAIN (adquirir, in this case) to be an EVENT and its RANGE (Grupo-Roche) to be HUMAN.[12] A THEME slot REQUIRES an EVENT for the DOMAIN (adquirir) and any OBJECT or EVENT for its RANGE (Dr-Andrew). These constraints are always very general, but still can help eliminate wrong attachments and word meanings.

4. The lexicon entry explicitly includes constraints that override or add to the above ontological constraints. Section 4.2 gave an example of when this is necessary. In our example, however, the two word-senses for ''adquirir'' have no explicit constraints in their lexicon entries.

5. Other structures in the sentence that are not explicitly specified by the lexicon entry can nonetheless modify the word in question. For instance, adjectives and PPs typically add slots to the TMR corresponding to the word they modify, even though they rarely are included in its lexicon entry explicitly. In this case, ''adquirir'' is modified by ''a-traves-de,'' which, depending on the meaning used, will either add a LOCATION slot or an INSTRUMENT slot to the TMR resulting from adquirir's analysis. In both cases, the slot will be filled by the TMR that results from ''compania,''[13] which maps into either a CORPORATION or a SOCIAL-EVENT (as in ''companionship''). The only interesting constraints that arise out of these combinations is that for the LOCATION meaning of ''a-traves-de,'' the DOMAIN (adquirir) must be a PHYSICAL-OBJECT (which it is not), whereas the INSTRUMENT meaning requires an EVENT. Although the LOCATION meaning of ''a-traves-de'' can be eliminated using these constraints, it does not help to further disambiguate ''adquirir.''

---

[11]It is clear that ORGANIZATION as AGENT or THEME should select different types of EVENTS than, say, HUMAN. As the uK ontology is refined, such knowledge will be added.

[12]Please see the discussion of metonymy in the next section to understand how ORGANIZATION (grupo-roche's meaning) can meet a HUMAN constraint.

[13]The lexicon entries for ''a-traves-de'' needs to be consulted to determine these facts.

8

5.2 Applying Constraints

uK employs an ontological graph search function (Onyshkevych 1995) to check constraints. This function determines relevant paths between two concepts and returns a score based on their degree of closeness. For example, **check-onto-con**(ACQUIRE EVENT)[14] returns a score of 1.0 (out of 1.0) since ACQUIRE is a type of EVENT. However, **check-onto-con**(ORGANIZATION HUMAN) returns a score of 0.9 along with the path (ORGANIZATION HAS-MEMBER HUMAN). This indicates that ORGANIZATION can stand in the place of HUMAN because it has HUMAN members. This and other types of metonymy are frequent in natural language and are detected automatically by uK.

5.3 Determining the Best Combination of Word Senses

The early versions of the uK analyzer at this point simply tried all of the possible combinations of word senses. Each combination activates the applicable constraints, which are combined into a total score for the combination. The combination with the best total score is chosen as the basic Semantic Dependency Analysis, the core TMRs to which other microtheories (such as aspect and coreference) can be applied. In the example sentence, the following choices were made:

1. ''a-traves-de'' is INSTRUMENT, since its LOCATION meaning would require ''adquirir'' to be a PHYSICAL-OBJECT.

2. ''en'' is LOCATION, since its TEMPORAL meaning requires ''espana'' to be a TEMPORAL-OBJECT.

3. ''adquirir'' maps into ACQUIRE, since its LEARN sense requires ''Dr-Andrew'' to be INFORMATION.

4. ''Dr-Andrew'' is an ORGANIZATION, since its HUMAN meaning cannot be the THEME of an ACQUIRE concept.

5. uK currently has trouble choosing between the CORPORATION and SOCIAL-EVENT meaning of ''compania,'' the object of the ''a-traves-de'' PP. Both can have locations in Spain, and both can be INSTRUMENTS of EVENTs. At this point, uK needs to add information into the ontology that ORGANIZATIONS can typically fill the INSTRUMENT slot of ACQUIRE acts, but SOCIAL-EVENTS cannot.[15] Statistical information[16] could also be consulted to tell us that in this business context the CORPORATION meaning is more likely.

5.4 Advanced Computational Methods

The uK project is experimenting with a number of computational techniques which aim to make it an efficient processor and, perhaps more importantly, a robust one which can handle a wide variety of input language, even language not specifically anticipated in the lexicon.

In addition to the microtheories that will be developed in the coming months to address specific language problems,[17] the analyzer utilizes an opportunistic, ''bulletin-board'' processing scheme which takes advantage of the following computational techniques:

1. Dependency Analysis. The key difficulty in natural language processing is the complex interplay of constraints present in even the simplest texts. Choosing one particular sense of a word may seem locally optimal, but it may create problems elsewhere which may ultimately lead to failure. In fact, in a typical problem, there are chains of dependency, where one choice eliminates choices at other points, which in turn eliminates other choices, etc.. The solution to these difficulties is a dependency-directed analysis which systematically tracks dependencies (Beale 94, Beale and Nirenburg 95) and can 1) propagate related constraints forward automatically, 2) automatically detect inconsistent solutions, and 3) be used in failure processing to determine the cause of failures and suggest recoveries.

---

[14] Which asks ''Is ACQUIRE an EVENT?''

[15] It is these type 2 constraints that the ontology needs updating the most.

[16] Statistical information is usually relegated to the role of optimization only, as described below.

[17] Microtheories of coreference, ellipsis, metaphor, time and aspect will be developed, among others.

2. ''Best First'' Processing. uK uses statistical data to determine the most likely senses of the input words. These senses are tested first, and if a result that ''satisfices'' is obtained, processing ends. This ''best first'' approach is extended to every aspect of processing, including failure recovery and ambiguity resolution.

3. ''Failure Recovery'' Techniques. Failures can arise from various sources. The actual input text can contain spelling errors. The syntactic analysis which is the input to uK can be in error. The lexicon and/or ontology can be erroneous or lack needed information. The analyzer itself can make incorrect decisions. uK tries to deal with these problems by:

    - using the dependency analysis to see why failures occurred
    - checking for metonymic/metaphoric language
    - if missing slot fillers, positing gaps (ellipsis)
    - changing syntactic analysis, including trying different attachments
    - relaxing thresholds
    - ordering possible recoveries using a sophisticated ''best first'' approach.

4. Ambiguity Resolution. If the basic semantic constraints cannot fully disambiguate, uK can:

    - use collocational preferences stored in the lexicon
    - use statistical methods to determine the most likely meanings
    - allow the ambiguity to remain. Subsequent clauses combined with coreferences might resolve the problem.
    - apply attachment rules such as ''referential success'' and/or ''minimal attachment''
    - Use ''expectations'' to moderate. For instance, in the current example, if one of the ''adquirir'' senses expected[18] an INSTRUMENT slot (which ''a-traves-de'' adds), favor that sense.

6 Conclusion

The Mikrokosmos Project at NMSU is the first, large-scale attempt at a knowledge-based machine translation system. We have successfully implemented the first and central stage of basic Semantic-Dependency-Structure building. This involved the creation of a large, language independent ontology which interacts with the Spanish semantic lexicon. The uK analyzer extracts semantic constraints from these two sources, analyzes them using a sophisticated ontological graph search function, and determines the combination of choices that yields the best overall score. Along the way, it utilizes the computational techniques described as dependency-directed processing, best-first processing, failure recovery and ambiguity resolution to ensure efficient and robust analysis.

In the coming year, uK will implement various microtheories which will build on the basic Semantic Dependency analysis. Included in this effort will be microtheories of ellipsis, metaphor, coreference and time and aspect. uK is also seeking to expand its language coverage. We hold great interest in the languages of South-East Asia, particularly Thai. For this reason we are especially grateful for the opportunity to present our work at this year's Symposium for Natural Language Processing in Bangkok.

References

Beale, S. and Nirenburg, S. (1995). Dependency-Directed Text Planning. To appear in Proc. IJCAI-95 Workshop on Multilingual Text Generation. Montreal, Canada.

Beale, S. (1994). Dependency-Directed Text Generation. Technical Report MCCS-94-272, Computing Research Lab, New Mexico State University, Las Cruces, NM.

Carlson, L. and Nirenburg, S. (1990). World Modeling for NLP. Technical Report CMU-CMT-90-121, Center for Machine Translation, Carnegie Mellon University, Pittsburgh, PA.

---

[18]''Expect'' meaning here that it is a slot explicitly defined for the concept.

Mahesh, K. and Nirenburg, S. (1995). A Situated Ontology for Practical NLP. To appear in Proc. IJCAI-95 Workshop on Basic Ontological Issues in Knowledge Sharing. Montreal, Canada.

Nirenburg, S. and Levin, L. (1992). Syntax-driven and ontology-driven lexical semantics. In *Lexical semantics and knowledge representation,* Pustejovsky, J. and Bergler, S. Eds. Heidelberg: Springer Verlag.

Onyshkevych, B. (1995). A Generalized Lexical-Semantics-Driven Approach to Semantic Analysis. Dissertation Proposal, Carnegie Mellon University, Program in Computational Linguistics.

Onyshkevych, B. and Nirenburg, S. (1994). The lexicon in the scheme of KBMT things. Technical Report MCCS-94-277, Computing Research Laboratory, New Mexico State University.

Viegas, E. and Nirenburg, S. (1995). The Semantic Recovery of Event Ellipsis: its Computational Treatment. To appear in Proc. IJCAI-95 Workshop on Context in NLP. Montreal, Canada.