Developing Knowledge Bases for MT with Linguistically Motivated Quality-Based Learning

Evelyne Viegas
New Mexico State University
Computing Research Laboratory
Las Cruces, NM 88003
USA
viegas@crl.nmsu.edu

Abstract

In this paper we present a proposal to help bypass the bottleneck of knowledge-based systems working under the assumption that the knowledge sources are complete. We show how to create, on the fly, new lexicon entries using lexico-semantic rules and how to create new concepts for unknown words, investigating a new linguistically-motivated model to trigger concepts in context.

1 Introduction

When parsing or generating, the MT system heavily depends on the completeness of the static sources available to it, in terms of Lexical Knowledge Base (LKB), Conceptual Knowledge Base (CKB) or ontology, and the grammar. Fighting the incompleteness of the static sources has been addressed by both the Computational Linguistic (CL) and the Artificial Intelligence (AI) communities. CL researchers extracted knowledge (usually lexical and/or syntactic) from *very large corpora.* This lexical acquisition has been mainly concerned with surface level co-occurring data (e.g., Zernik and Jacobs, 1990), hyponymy extraction based on syntactic criteria (e.g., Hearst, 1992) or lexico-semantic associations (e.g., Resnik, 1992). However these proposals cannot go beyond a shallow level of learning (selectional restrictions). Yet understanding needs a more fine-grained conceptual knowledge, in terms of roles and role filler constraints (e.g., Mahesh *et al.* (1997a) for a discussion on the necessity of such knowledge.) In this proposal, we only use *positive occurrences of a small corpus of one text* (there is however no limitation on the length of a text; in fact, the longer, the better) to avoid dealing with multiple meanings of a word used in different contexts (e.g. homonymy). AI researchers used standard machine learning algorithms or adapted them to the language understanding task (see Wermter *et al.* (1996), for a review). As pointed out by Schnattinger and Hahn (1998), the main drawback of these proposals is that they keep the understanding and learning modes separate. However the interpretation of a text depends on both, meaning that these modes have to interact with each other. In this proposal, we advocate a *knowledge-intensive model of word and concept learning* which interacts with the non-learning mode of text understanding.

At the lexical level (Section 2), we make use of lexico-semantic rules (LSRs) to create new entries. LSRs create new syntactic and semantic mappings in a new entry. At the conceptual level (Section 3), we investigate a model to create new concepts for unknown words. To do so, we use the analysis of the text to provide the unknown word with its conceptual environment in a given context, and a concept trigger model, we call **Concept Trigger** (CT), which builds the semantic space around the new concept (in other words, the conceptual frame) and attempts to place it in the CKB or ontology hierarchy. Briefly, when a word in a text is unknown to the system, i.e. it does not belong to the lexicon, lexico-semantic rules cannot be applied to the root of the word (i.e. after the word has been lemmatized), then CT is activated to create a semantic mapping for the word. To do so, various data are consulted at run time by CT, namely:

- the existent CKB which includes domain knowledge,
- the semantic representation of the text being processed providing the unknown word with a contextual environment, and

- a database of linguistic operators, discussed in Section 4, which helps establish the relation between the unknown word and the co-occurring words, thus enabling CT in making hypotheses on the semantic type of the unknown word.

This research has been carried out within the KB paradigm of Mikrokosmos (Nirenburg *et al.*, 1996).[1] This is not because such a system is a prerequisite to the approach, but rather because we are intimately familiar with it, rendering experimentation easier to perform and also because it already rests on the principles we wish to validate (e.g. use of KB systems for high-quality MT). In this paper, we do not discuss general principles. We focus on the automatic extension of LKBs and CKBs.

2 Lexico-semantic Rules and New Lexicon Entries

Lexicon entries are mapped to a CKB or ontology. The CKB is a large collection of information about EVENTS, OBJECTS and PROPERTYS in a domain. In Mikrokosmos, the ontology consists of concepts (named sets of property-value pairs) organized heterarchically on subsumption links, with about 14 links between concepts (e.g. AGENTOF; INSTRUMENT; EFFECT; INTERNATIONALATTRIBUTE; see also the relations in TEACH below) (Mahesh, 1996).

Figure 1 illustrates relevant aspects of the syntax, semantics and linking information of a lexicon entry of the English word *teach*.[2]

The feature "sem" gives the semantics of the English word *teach* for its sense TEACH. This latter is a well-defined symbol or concept in the Mikrokosmos ontology as described in Mahesh (1996). We present below an example of conceptual information found in the ontology for TEACH.

```
[key: "Teach",
def:[text:"To explain or teach someone about some
        object, subject matter, or concept."]
sem:[name: Teach, relations:
```

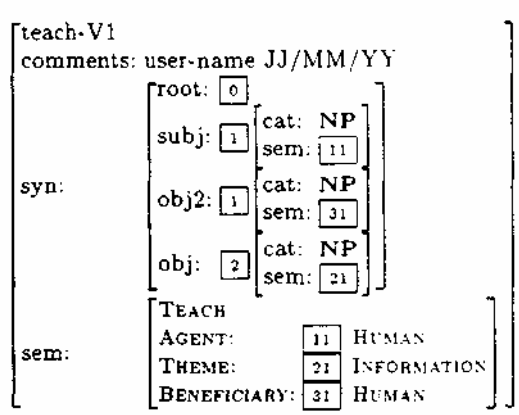

Figure 1: Partial Sense Entry for *teach*.

```
        < % A list of relations %
[name:Isa,
 dom:<AcademicActivity CommunicativeEvent>] ,
[name:Agent,
 dom:<Human>,
      default:<AcademicSpecialist Teacher>],
[name: Theme,
 dom: <Information>],
[name: Effect,
 dom: <Learn Understand>],
[name: Intention,
 dom: <Education>],
[name: Experiencer,
 dom: <Animal>, default: <Student>],
[name: HasAudience,
 dom: <Human>, default: <Audience>],
[name: AreaOfActivity,
      dom: <FieldOfStudy>],>]]
```

Acquiring the sense feature "sem" represents the most expensive operation in the acquisition of computational semantic lexicons, as the task involves human judgments and cannot be completely automated. We argue that once a core lexicon has been acquired (about 7000 sense entries) then one can automatically extend this core lexicon providing a coverage good enough • to process a text.

In the following, we show the advantage of using LSRs. First, they can be considered as a means to reduce the number of lexicon en-

[1] We do not have room here to describe Mikrokosmos. We refer the reviewer to papers in http://crl.nmsu.edu/Research/Projects/mikro/index.html

[2] The syntactic information is not complete here. Rules (e.g. alternation rules) can be applied to this base form to produce new subcategorizations (e.g. subj-obj-[to]obj2 for *teach).*

try types, and generally to make the acquisition process faster and cheaper. Second, they can enhance the results of analysis processing by creating new entries for unknown words from the lexicon found in corpora. Lexical rules have been addressed by many researchers, e.g. Leech (1981), Ostler and Atkins (1992), Copestake and Briscoe (1996), and others. More specifically, Onyshkevych (1999) addresses the theoretical background of Lexical Rules (LRs) mentioning three different types of LRs: i) inflected forms (passivation - dative alternation); ii) word formation (derivational morphology) and iii) polysemy (sense extension - type coercion). It also addresses the pros and cons on when to apply the rules (acquisition time - lexicon load time or run time). Note that hybrid scenarios are also plausible: for example, LRs can be applied at acquisition time to produce new lexical entries, and may also be available at run time as an error recovery procedure to attempt generation of a form or word sense not in the lexicon, as done for instance in Viegas *et al.* (1996a).

To sketch this operation briefly, applying LSRs to the entry for the Spanish verb *comprar* (buy) produced automatically 26 new entries *(comprador-N1* (buyer). *comprable-Adj* (buyable), etc). This includes creating new syntax, semantics and syntax-semantic mappings with correct subcategorizations and also the right semantics. For instance, the lexical entry for *comprable* will have the subcategorization for predicative and attributive adjectives and the semantics adds the attribute FEASIBILITYATTRIBUTE to the basic meaning BUY of *comprar.* The form list generated by the morpho-semantic generator (e.g. *comprable, comprador, recompra,* etc.*), is* checked against Machine Readable Dictionaries (MRDs) and dictionaries and the forms found in them are submitted to the acquisition process. However, forms not found in the dictionaries axe not discarded outright because the MRDs cannot be assumed to be complete. Some of these "rejected" forms can, in fact, be found in corpora or in the input text of an application system.

Viegas *et al.* (1996b) describes about 100 LSRs which were applied to 1056 verb citation forms with 1,263 senses among them. The rules helped acquire an average of 26 candidate new entries per verb sense. This produced a total of 31,680 candidate entries, with an average of over 90% and 85% correctness in the assignment of syntax and semantics respectively.

LSRs are composed of lexical semantics, where the left-hand side of the rule specifies which information to look for in the entry of the dictionary and the right-hand side specifies the information which should be put in the new entry. They are also linked to morpho-semantics, with the rule specifying which semantics are attached to which affix (e.g., *able* can be attached to any EVENT which has both an AGENT and a THEME).

Here we use the same mechanism to enhance the results of analysis processing by creating new entries for unknown words found in corpora, on the fly.

Formally, LSRs are rules. Practically, they create new entries from existing entries in the lexicon that they modify. We discuss below an example involving the *LSR2AgentOf* rule.

For instance, *LSR2AgentOf* is a rule which takes as input a verb and modifies its POS, syntax and semantics. The "derived" entry has N for the POS, the subcategorization of a noun for the syntax and add AGENTOF to the event of the old entry. For instance, let us consider the entry *drive* mapped to the concept DRIVE in the ontology (Figure 2). The selectional restrictions of *drive* are further constrained to HUMAN (with a Default DRIVER) for AGENT and PHYSICALOBJECT (with a Default WHEELEDENGINEVEHICLE) for THEME.

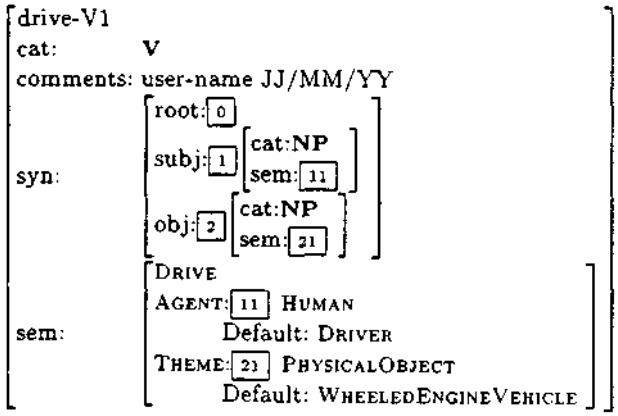

Figure 2: Partial Entry for the English word *drive.*

If we apply the LSR *LSR2AgentOf* to the

entry *drive,* it will produce the new entry for *driver* as illustrated in Figure 3.

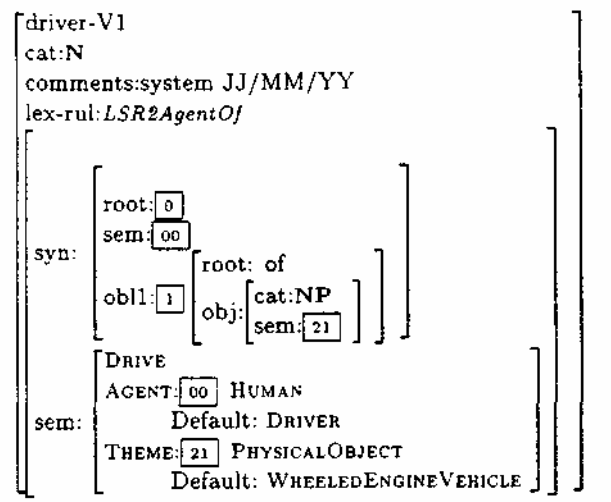

Figure 3: Partial Entry for the English word *driver.*

The entry for *driver* specifies the new POS, a new "SYN" with an optional prepositional phrase (opt +), and a semantics which is the same as the one for *drive.*

LSRs (such as *LSR2AgentOf)* are language independent, and can therefore be used for the production of new entries for any language, as far as the semantics is concerned. It is also regular as far as POS and syntax is concerned for family-related languages.[3] The morpho-semantic generator is obviously more language dependent as to which affix is assigned which rule is language dependent. For instance, in Spanish *-dor is* assigned *LSR2AgentOf,* whereas *-er* and *-eur* will be assigned the same rule for English and French, respectively. Viegas *et al.* (1996b) showed ways to avoid overgenerations of forms in the case of Spanish, which is more productive than English from a morphological viewpoint. Ultimately, the lexicon entry will have to be checked by a human to prune word forms with a wrong semantic assignment. For example *revolver,* which would be created from *revolve* with the semantics of *LSR2AgentOf,* thus would be missing the WEAPON meaning. However, in this research, we are using the LSR for analysis. As such, overgeneration of forms is not a problem as one does not expect to find a wrong form in a corpus. Overgeneration of forms appears in cases of "blocking," as identified by Briscoe *et al.* (1995). For instance, the system will successfully apply the *LSR2AgentOf* to the entry *steal,* producing *stealer,* whereas this form should be blocked by *thief* in the lexicon. It is important to filter out these "wrong" forms for generation (by trying to match the created form against an on-line dictionary for instance). But when using the LSRs at run time and for analysis, there is no need for such checking.

Moreover, in the approach proposed here, it is still possible to ignore some of the "incorrect" entries using the selectional restrictions of the other words with which the entry created automatically co-occurs in context. In other words, we evaluate the "correctness" of the entry in the light of the other words in which it appears in the text. For instance, in *He killed Paul with this revolver,* if the system can use the selectional restrictions of *kill* such as (AGENT: ANIMAL, THEME: ANIMAL, INSTRUMENT: ARTIFACT), then *revolver as* the HUMAN AGENTOF a "revolving" event will be dismissed, at the benefit of ARTIFACT - INSTRUMENTOF KILL. This is one of the reasons why the learning mode must interact with the parsing mode: context helps build hypotheses on the validity of the new entry.

When multiple LSRs can be applied to an unknown word, all are tried and the semantic constraints of the co-occurring words provided by the analysis parser (e.g. the Text Meaning Representation (TMR) in Mikrokosmos) filter out the ones which do not fit, as described in Beale *et al.* (1995). Mahesh *et al.* (1997b) showed that the quality of the TMR (on the word sense disambiguation task) did not fully degrade with unknown words. Therefore, here we start the interaction between the understanding and learning modes on an already built TMR to restrain the search space of hypotheses. We present an extract of the TMR for (1) below assuming *teacher* has not been found in the LKB.

(1) *"There is incredible pressure on school systems and* **teachers** *to raise test scores," says...,* [Wall Street Journal].

[3] Note that for unrelated languages, the POS and syntax will necessitate a thorough manual checking.

```
PROPOSITION-1
RAISE-1
AGENT: teacher-1*unknown*
THEME: SCORE-1

SCORE-1
PURPOSE: TEST-1
```

Here, *teacher* will be processed by a morphological analyzer producing the root *teach* (which we assume here to be available in the lexicon), and the suffix *-er,* which is available in the data bank of affixes and will be assigned the rule *LSR2AgentOf* as follows:

teacher: TEACH, *-er LSR2AgentOf*

However, *teacher* is a social category as well as a *person who teaches;* its lexicon entry (Figure 4), which has been automatically created, accounts for the latter but not the former. In other words, LSRs help construct the semantics HUMAN AGENTOF TEACH but not the social role TEACHER which should be in the subtree of ACADEMICROLE.

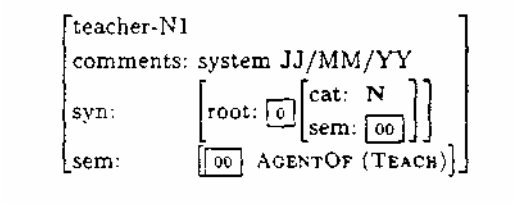

Figure 4: Partial Sense Entry for *teacher* Derived from *teach* with *LSR2AgentOf.*

Accounting for the former involves two options:

1. The concept TEACHER already exists in the ontology under, say, the concept ACADEMICROLE.
2. The concept TEACHER does not exist in the ontology and we want to create it and automatically insert it in the CKB. Here we have two subcases:
   - 2.a the entry for *teacher was* created by an LSR;
   - 2.b no LSR could be applied to create an entry for *teacher.*

Case 1. is easy as we can directly replace AGENTOF TEACH by TEACHER. This information is already part of the frame of TEACHER in the CKB.

Case 2. is more complex. Briefly, the idea is to create a new concept, labeled as the lexical item preceded by TC_. We use this notation for convenience purposes only, to help the CKB developer to semi-automatically check the new concept created at run time during maintenance checking. In 2.a, the idea is to introduce the new concept in the CKB. The new concept TC_TEACHER will have minimally encoded that it is AGENTOF TEACH as part of its frame in the CKB. In 2.b, no LSR could produce information and here the system primarily relies on the semantic information provided by the semantic analyzer on the words co-occurring with the unknown word (e.g., in the TMR).

In this section, we dealt with the incompleteness of the lexicon and ways to fight it by applying LSRs on the lemmas of unknown words. This resulted in two possibilities: a) a rule (or more than one) could successfully be applied, then producing an entry (multiple entries) for the unknown word with POS, syntactic information and semantic information; or b) no rule could be applied, in which case the word was given the label of the unknown word preceded by CT_.

In next section, we investigate a new model to create a new concept using the existent CKB and the semantic context in which the unknown word appears. The model helps in solving 2.a and helps in making hypotheses in the case of 2.b.

3 The Concept Trigger Model

A concept from the CKB, once acquired or checked by the CKB developer at acquisition or maintenance time, is considered a "constructed" concept. Otherwise it is a "triggered concept" (TC), automatically created at run time by CT and waiting for manual checking to get (or not in case it is wrong) the status of "constructed concept" and become an integral part of the CKB. The manual checking of the TC consists in verifying its place in the conceptual hierarchy, i.e. verifying its ISA and SUBCLASS links, and then checking the inherited

links and eventually adding more roles (e.g. in TEACHER, one could add PURPOSE: EDUCATE) and defaults to inherited roles (e.g., LOCATION: PLACE; DEFAULT: CLASSROOM).

CT uses four resources: i) a CKB; ii) an LKB (e.g. lexicons developed for Mikrokosmos in Spanish, English and Chinese, with 35,000, 12,000 and 2,500 word sense entries respectively); iii) the meaning of the text (expressed with concepts from the CKB) where the unknown word appears (e.g. TMR in Mikrokosmos); and iv) linguistic operators helping identify semantic relationships between concepts, as discussed further.

To illustrate iii), let's assume that "revolver" is unknown in (2), then the selectional restrictions of *hand-out* will help create the hypothesis that the THEME must be of type OBJECT.

(2) *At a recent dinner for institutional investors, a New York portfolio manager says, the mood was so bleak that "if they had handed out* **revolvers** *instead of cigars, some people would have shot themselves."* [Wall Street Journal]

A further hypothesis, derived from the conceptual information of *shoot oneself* mapped to the concept KILL can help derive the hypothesis that it could be of type DEVICE (ISA ARTIFACT) via the INSTRUMENT filler of KILL. This real life example shows the complexity of deriving the right hypotheses. We cannot yet show how to derive the right relation, rather we aim at providing the system with the set of valid possibilities, going as far down the hierarchy of concepts in the CKB as possible.

3.1 Towards a Characterization of CT

A concept cannot be acquired independently from other concepts. **CT** stands for that purpose, where a concept is being related to the other concepts triggered from the analysis of the text and the CKB. **CT** helps build the conceptual frame of the new concept.

CT relies on different kinds of knowledge used to grasp the semantics of the TC. We illustrate through TC_TEACHER associated to the unknown English word *teacher* which kind of information enters in a TC:

1. The "Natural Logic" zone contains information concerning the semantic space of the TC, by means of type/sub-type relations, including all the role fillers available to the immediate hyper-/hyponyms, (e.g., ACADEMICROLE, UNIVERSITYFACULTY, ACADEMICBUILDING, ...);
2. The "Encyclopedic" zone provides information on encyclopedic knowledge, providing the TC with its associated concepts, for a particular domain: for instance in the academic domain we will find at least SUBJECT, COURSENUMBER, ... This information is subdivided in three sub-zones, namely intrinsic, extrinsic and functional.
3. The "Lexical" zone points to an entry in the lexicon, when created by an LSR.

For instance in TC_TEACHER one could find the following:

```
[Tc_Teacher(X):
  [natural_logic:
    Teacher(X),Professor(X),Lecturer(X), ...]
  [encyclopedic:
  [intrinsic: Name(B), Address(C), ...
  extrinsic: Student(Y), ...
  functional:Teach(X,Subject,Student), ...]]
  [lexical: teacher(X)]]
```

To "trigger" knowledge in the TC we rely on the KBs and the TMR. In the case of *teacher,* in 2.a, the system has the knowledge that it is AGENTOF TEACH: **CT** retrieves all AGENTOF TEACH from the CKB, constructing the "natural logic zone." We also rely on linguistic knowledge from the text where the system labels the different relations found between the unknown word and its co-occurring words. To do that, we use linguistic operators, modeled to provide interconceptual relations between the Tc and other concepts in the semantic representation of the text (e.g., TMR), as described in next section.

3.2 From Linguistic Operators to Interconceptual Relations

From a linguistic point of view, in the case of the "logical zone," the interconceptual relations belong to the linguistic domain of BE *(teacher/professor/...).* In the case of the encyclopedic zone, the interconceptual relations belong to the linguistic domains of HAVE and

DO *(teacher/students)*. Here the operator can be lexicalized into a static verb *(my teacher* has *60 students),* or into a dynamic verb *(the teacher* **lectures** *students)*. Linguistic operators are mapped to interconceptual relations, which capture generalizations across languages. Linguistic operators and interconceptual relations help build hypotheses gathered in TCs, using a quality-based learning approach as described in Schnattinger and Hahn (1998) for instance. We base our qualitative ranking of the linguistic operators on linguistic constructions AND available semantic information coming from the existent CKB. The ranking follows a partial order >. In the case of some constructions into which a noun can enter, we have the following order:

N V N (clause - V-DO)) N's N (possessive)
N's N } N of N (possessive, membership - of-HAVE or descriptive - of-BE-IN)
N's N) N have N
N have N) N be N
N be N } NN (compounds)[4]

We have discovered three operators for BE and HAVE **(Identification (Iden), Differentiation (Diff) and Localization (Loca))**, which can further be subdivided along intrinsic (intr) or extrinsic (extr) properties, as illustrated below through examples:

> $Iden_{intr}$: he **IS** a teacher; the city **OF** Singapore.
>
> $Iden_{extr}$: she **IS** nice.
>
> $Diff_{intr}$: he **HAS** blue eyes; **HER** green eyes; he **IS** blue-ey**ED**; the legs **OF** the table.
>
> $Diff_{extr}$: she **HAS** a car; john'**S** car; the car **OF** the company.
>
> $Loca_{intr}$: the City **OF** London.
>
> $Loca_{extr}$: he **STANDS IN** the garden.

For instance, "OF" and "ADJ-N_ed" (blue-eyed) can be linguistic traces of HAVE. Depending on the semantics of the nouns, "of" will be further categorized as $Diff_{intr}$ or $Diff_{extr}$. These operators are organized into a hierarchy. For instance $Diff_{intr}$ and $Diff_{extr}$ are subtypes of *Diff;* the more semantics available, the more distinctions can be made in the operators, the higher is the score attributed to the operator.

Any relation belonging to BE can be mapped to either $Iden_{intr}$ or $Iden_{extr}$ under the following conditions:

$Ident_{intr}$: $(x_{pos} : N, x_{sem} : Human, y_{pos} : N, y_{sem} : SocialRole \rightarrow HumanSocialRoleRel(y,x))$

This means that if the first noun (x) is of type HUMAN and there is a HUMANSOCIALROLEREL between x and y (also a noun), then y must be of type SOCIALROLE.

$Ident_{extr}$: $(x_{pos} : N, x_{sem} : Object, y_{pos} : Adj, y_{sem} : Relation \vee Attribute)) \rightarrow Property\ (x,y))$

With $Ident_{extr}$, if x is a noun of type OBJECT and y an adjective and there is a PROPERTY between x and y, then y must be either a RELATION or an ATTRIBUTE.

For instance, the operators between a TC and the concepts of the "natural logic" zone belonging to the domain of BE can be an **Iden** or a **Loca** interconceptual relation. The intrinsic and extrinsic sub-zones of the "encyclopedic zone" belong to the domain of HAVE and are a **Diff** relation.[5]

The TC_TEACHER type-example below provides an illustration of the interconceptual relations:

```
[tc_teacher(X,Y):
  [iden: professor(X), lecturer(X),...
    diff: name(B), student(Y),...
   oper: teach(X,subject,student),...]]
```

To make assumptions on the hypotheses, we follow the proposal of Schnattinger and Hahn (1998). However, developing linguistic operators and interconceptual relations enables us to get a better set of accurate hypotheses than relying on linguistic constructions alone as done by Schnattinger and Hahn. On the experiments done, we noticed that a finer-grained ontology and finer-grained linguistic operators reduce the number of hypotheses.

Then inserting the concept TEACHER at the right place in the hierarchy, is more difficult. It consists in finding a frame (from the

[4]See Viegas *et al.* (1999) for the description of the semantics involved in compounds.

[5]The operator DO is still under study.

frames associated to the concepts of the "natural logic" zone), which has the closest constraints compared to the partial frame associated to TEACHER (namely, AGENTOF TEACH). This new frame will have to be checked by a human ultimately.

4 Perspectives

In this paper, we have outlined solutions and directions to bypass the bottleneck of KB systems by providing ways of expanding the knowledge sources (LKB and CKB) at run time. We believe that this constitutes a necessary step in the development of future knowledge-based high-quality MT systems.

We have used LSRs to create new lexical entries, and investigated a new model called **CT** to create new concepts for unknown words. **CT** makes use of conceptual knowledge from the existent CKB and also of linguistic operators to help identify interconceptual relations and concepts in context. **CT** interacts with the non learning mode of the analyzer to make hypotheses on the semantics of the unknown word. At this stage, the automatically created entries can help in better disambiguating a text by having **CT** feed the analyzer with valid hypotheses on the semantics of the unknown word. The issue of placing the new concept at the right place in the conceptual hierarchy is more complex and still under study. The proposal developed in this paper helps in recovering or repairing from incomplete inputs to the analyzer. Therefore, putting the concept at the right semantic sub-tree (e.g., TEACHER placed in SOCIALROLE or ACADEMLCROLE and not say under EVENT) seems reasonable.

Further research concerns applying the same methodology to more unknown words per text and also among interdependent unknown words per text. If the unknown words do not interact, then we expect that the same approach be used; if the unknown words are connected, hypotheses made by **CT** should have a lower confidence than when only one word is unknown.

We have starting exploring this framework on Spanish text too, showing that our methodology is language independent. Linguistic operators are also language independent. What is language dependent is the realization of the operators.

Yet another angle of this research concerns using **CT** for large-scale acquisition of concepts and therefore developing the algorithm to place the concept at the right place in the conceptual hierarchy. The best way to do so might be to move towards an interactive system of acquisition of concepts where a human immediately revises the suggestions made by **CT** and **CT** learns from human input.

Acknowledgments:

This work has been supported in part by DoD under contract number MDA-904-92-C-5189. We are thankful to the anonymous reviewers.

References

Beale, S., Nirenburg, S. and K. Mahesh (1995) Semantic Analysis in the Mikrokosmos Machine Translation Project. In *Proceedings of the 2nd Symposium on Natural Language Processing,* Bangkok, Thailand.

Briscoe, T., Copestake, A. and A. Lascarides (1995) Blocking. In P. Saint-Dizier and E. Viegas (eds.) *Computational Lexical Semantics.* CUP.

Copestake, A. and T. Briscoe. 1996. Semi-Productive Polysemy and Sense Extension. In *Journal of Semantics, vol.12.*

Hearst, G. (1992) Automatic acquisition of hyponyms from large text corpora. In proceedings of *Coling92,* Vol 2.

G. Leech. 1981. *Semantics.* Cambridge: Cambridge University Press.

Mahesh, K. (1996) *Ontology Development: Ideology and Methodology.* Technical Report MCCS-96-292, CRL, New Mexico State University.

Mahesh, K., Nirenburg, S. and S. Beale (1997a) If you have it Flaunt it. In Proceedings of *TMI97,* Santa Fe.

Mahesh, K., Nirenburg, S., Beale, S., Viegas, E., Raskin, V. and B. Onyshkevych (1997b) Word sense disambiguation: why statistics when we have these numbers? In Proceedings of the *7th International Conference on Theoretical and Methodological Issues in Machine Translation,* Santa Fe, NM, 151-159.

Nirenburg, S., Beale, S., Helmreich, S., Mahesh, K., Viegas, E. and R. Zajac (1996) Two principles and six techniques for rapid MT development. In Proceedings of *AMTA96.*

Onyshkevych, B. (1999) Categorization of Types and Application of Lexical Rules. In E. Viegas (ed.) *Breadth and Depth of Semantic Lexicons.* Dordrecht: Kluwer Academic Press.

Ostler, N. and S. Atkins. 1992. Predictable meaning shift: Some linguistic properties of lexical implication rules. In J. Pustejovsky and S. Bergler (eds), *Lexical Semantics and Knowledge Representation.* Berlin: Springer, 87-100.

Resnik, P. (1992) A class-based approach to lexical discovery. In the proceedings of *ACL92.*

Schnattinger, K. and U. Hahn (1998) Quality-Based Learning. In Proceedings of the *13th European Conference on Artificial Intelligence.*

Viegas, E., Onyshkevych, B., Raskin, V. and S. Nirenburg. (1996a) From *Submit* to *Submitted* via *Submission:* on Lexical Rules in Large-scale Lexicon Acquisition. In *Proceedings of the 34th Annual Meeting of the Association for Computational Linguistics,* Santa-Cruz, CA.

Viegas, E., Gonzalez, M. and J. Longwell. (1996b) *Morpho-semantics and Constructive Derivational Morphology: a Transcategorial Approach to Lexical Rules.* Technical Report MCCS-96-295, CRL, NMSU.

Viegas, E., W. Jin and S. Beale (1999) Long Time No See: Overt Semantics for Machine Translation. In Proceedings of the International Conference on *Theoretical and Methodological Issues in Machine Translation,* Chester, UK.

Wermter, S., Riloff, E. and G. Scheler (1996) Learning Approaches for Natural Language Processing. In Wermter, S., Riloff, E. and G. Scheler (eds.) *Connectionist, Statistical and Symbolic Approaches to Learning for Natural Language Processing.* Springer.

Zernik, U. and P. Jacobs (1990) Tagging for Learning: collecting thematic relations from corpus. In Proceedings of *Coling90,* Vol.1.