Interpreting Common Words in Context : a Symbolic Approach

Violaine Prince

LIMSI-CNRS.
P.O. Box 133 F91 403 Orsay cedex France
tel. : (33 - 1) 69 85 80 18
fax: (33-1) 69 85 80 88
e-mail : prince@limsi.fr

Content areas : Natural Language Processing; Semantic Interpretation.

Abstract

This paper presents a lexical model dedicated to the semantic representation and interpretation of individual words in unrestricted text, where sense discrimination is difficult to assess. We discuss the need of a lexicon including local inference mechanisms and cooperating with as many other knowledge sources (about syntax, semantics and pragmatics) as possible. We suggest a minimalist representation acting as a bridge between a conceptual representation and the microscopic sense variations of lexical semantics. We describe an interpretation method providing one or many alternative candidate(s) to the word, as representatives of its meaning in the sentence (and text). Our system, presently implemented and tested on a restricted scale, belongs to an existing multi-expert architecture for natural language processing, whose ambition is to provide a tool-box for applications involving language understanding.

1.Introduction

Intelligent lexica, i.e. the widespread desgination for lexical knowledge bases, are more and more frequent in computational natural language processing. There is an important need for sophisticated knowledge sources, because language understanding by computers evolves towards real-case applications. An intelligent lexicon could be seen as a source of knowledge about lexical items, not only accounting for hierarchical semantic relations (such as in semantic nets) or case relations (such as in conceptual graphs), but also ranging from morphological information to common usage knowledge. This is what makes an intelligent lexicon different from both a conceptual semantic representation (like both semantic nets and conceptual graphs) and a 'flat-entry' dictionary (i.e. with morphological indications and a literal sense only).

In this contribution we present a lexicon belonging to the category of intelligent lexica, which is specialised for the semantic interpretation of numerous 'common words'. Common or usual words tend to be the most difficult to interpret in terms of word sense discrimination, because there is a high probability to find them in many different contexts with many plausible interpretations. For instance, we would like to be able to interpret the word *father* in these sentences:

(1) The ***father*** *process forks and generates a son process (in Unix).*
(2) His ***father*** *is a weak and benevolent man.*
(3) Ritchie and Thompson were the ***fathers*** *of Unix.*
(4) Our ***fathers*** *used to dwell in dark caves.*

where very different contexts of use and of meaning are involved (technical context in sentences (1) and (3), more general context in sentences (2) and (4); metaphoric use in sentences (1) and (3), more general use in sentences (2) and (4)).

We argue about the necessity to have such a lexicon as a complementary knowledge source to other and more traditional semantic, syntactic and morphologic representations, belonging to an already existing AI architecture dedicated to natural language processing. We also show how it fits into the on-going research in symbolic AI about lexical knowledge representation, and try to demonstrate the utility of our approach.

2.Relating to other works and situating our approach

2.1 Relevant recent history of semantic lexical design

In the early eighties, works have shown that lexical semantics could not be approached by a schematised word-concept relation. The main milestones, in symbolic AI, were Small and Rieger's word-expert (1982),, the computational models based on the exhaustive survey of lexical semantic relations in (Evans *et al.* 1983), and the trend in lexical design that was following systemic grammars such as Cumming's master lexicon (1986). Many scientists demonstrated that the complexity of lexical semantics did not necessarily lead to such over-extensive representations of word senses, but could be adequately replaced by reasoning mechanisms. One of the first to open the way was Stallard (1987) who suggested a minimal core sense, associated with dynamic features, which would infer the possible relations (instead of representing them). Two families of recent works, favouring this approach, could be found. One is characterised by a complete but multi-faceted lexicon, associating descriptive minimal structures with inference rules of the 'coercion' type: a major representative is Pustejovsky's generative lexicon (1991), and so is the feature-based system with default rules proposed by Briscoe and Tappleton (1991). The main problem with these lexica remains the size versus completeness problem. To what extent one must account for existing or derivable senses ? The other family is composed of differentiated lexica, according to their use by the system. A typical and recent representative is McRoy's system (1992): it provides specialised and related lexica serving as input (or output) to expert modules in a word sense discrimination architecture. This system is quite interesting, but in the referenced contribution, only part of the possible interpretation problems has been dealt with and one does not see exactly how a complete text could be interpreted with this system alone.

In parallel with these results, interesting findings about lexical featuring have been achieved by people trying to solve non-literal senses in figures such as metonymy and metaphors. Martin (1988; 1990), when developing the Midas system, relied on the Kodiak introduction of a metaphorical relation between individual concepts (Wilensky 1986): a word sense could be addressed by a concept of a different designation. This aspect has scarcely been highlighted by traditional lexical design. However, the Midas system is dedicated, and doesn't seem to be able to cooperate with other knowledge sources and mechanisms. The same for Fass' Met* system (1991), which accounts for both metonymic and metaphoric senses. It is more complete, but still dedicated to non-literal senses caused by semantic or pragmatic aspects of language use. About lexical items, the author clearly indicates the existence of three lexical semantic constraints in contextual word sense discrimination: **preference** (based on Wilks' theory (1978)) which discriminates the semantic class of co-occuring items, **assertion** that imposes information to items posessing similar lexical syntactic constraints, and **relevance**. These aspects introduced the importance of context information, which wasn't widespread in lexical design.

2.2 Similarities and differences of our approach

Our problem was quite different from many other approaches to lexical design, in the sense that it had to fit into an existing architecture for natural language processing (NLP), described in (Sabah 1990), based on a blackboard structure and a control module handling specialised 'experts' , dedicated to specific subtasks of NLP such as : parsing, dialogue managing, reference resolution, ellipses resolution, topic determination, response planning, etc. Many aspects of ambiguity resolution were already achieved (focused on in (Prince and Sabah 1992)) by the following modules : a deterministic parser, with a coverage of 15 000 words, a topic determination module (which already solves problems dealing with sheer homonymy) a reference resolution module (for anaphors and ellipses), a spelling corrector, and most of all a conceptual semantic knowledge base, based on conceptual graphs modelling (Sowa 1984), comprising the semantics of most predicates (both verbs and nouns). As such, the semantic model was unable to discriminate very accurately word senses, and particularly not non-literal senses which were not recorded before. The architecture was lacking two modules : a system for accurate word sense discrimination, which is the subject of this contribution, and a system for metaphors and metonymy interpretation, which is anyhow linked with the first module.

Thus we had designed and implemented a lexical model with resolution mechanisms that satisfy the three following requirements : (i) account for the three lexical constraints described by Fass and Wilks; (ii) be as much as possible compatible with a conceptual graph semantic representation (at least at the heading level) ; (iii) provide a cooperating multi-expert architecture to be compatible with the existing NLP system. To these three, we have added two requirements : have the most compact representation ; do not try to solve sense discrimination like a yes-or-no case, because polysemy in common words is more a problem of how to separate slightly different shades of the same colour than a problem of how to choose between wholly distinct colours.

Hence, our semantic lexicon specialises in the representation and interpretation of words (mostly nominal) in unrestricted contexts, whose multiple senses are not easily discriminated, neither by function, nor by case structure, nor by topic determination, but which necessarily require many types of information prior to sense assignment. The descriptive structure is both minimalist and unique. However, effort has been put on reasoning mechanisms: inference rules are of three types. (i) Pragmatic rules which, conversely to all traditional methods, trigger sense instead of completing it, taking into account information similar to the three lexical semantic constraints in Fass and Wilks' theory, adding to them commonsense knowledge, generally absent from lexica. (ii) Semantic constraints handling semantic consistency in interpretation. (iii) Default reasoning, as in Briscoe's system, trying to derive information from a partially instanciated form in order to complete it. Like McRoy's, our lexicon is linked with other semantic knowledge bases provided by a natural language processing environnement: it is mainly related to a semantic network of ontologic (*is-a*) and meronymic (*part-of, member-of)* relations between concepts, and a conceptual graph base describing the case structures of verbal predicates. But unlike general-purpose lexica, our system has been able to link words and those concepts that are individual metaphors to these words, without the requirement of a specific metaphoric link as in Midas. Metaphoric sense is as 'natural' in our system as the literal sense.

Finally, the originality of our approach resides in the fact that links between words and concepts in our lexicon are only **relevance** links (Rieger 1984) : they impose no requirements on concepts linked to words, apart from their local description in terms of appropriate features. The relevance of a concept to a word sense is asserted by means of its contextual activity: if a concept is relevant to a sense in a given context, wether by affirmation or by negation, then it is a highly active partner in the word sense discrimination process; if a concept is implicitely pointed at, then it is considered as background information, not to be discarded. If it is irrelevant, then it has to be withdrawn.

3.Lexical model

The structured lexicon that we have designed is composed of the following items.

To every word w in the lexicon is associated a **descriptive structure** DS_w, which binds the word to concepts traditionally used to discriminate the word sense when this word is used in language. DS_w is a part of the semantic potential of the word, which we detail in § 3.1.

To every word w in the lexicon is associated a **set of rules** P_w, which launches the interpretation of that word in a sentence by relying on various information ranging from morphosyntactic knowledge to pragmatic knowledge. These rules are meant to indicate how relevant is a particular component of the word descriptive structure by 'measuring' its activity in the context of the sentence (respectively paragraph and text) where the word w has occurred.These rules are described in § 3.2.

To every word w is associated an **information domain** I_w , which stores the context knowledge necessary to launch word sense discrimination, by triggering rules of P_w. I_w is used in § 3.2.3.

All these elements are recorded for every lexical item belonging to the lexicon. The structured lexicon contains presently around 150 semantic structures (potentials) each being available for all the forms of a given lexeme, including gender, tense, plural, and adjective derivations (of common nouns or of verbs). We record only the root form of the word, leaving derivation links to an associated morphological dictionary of 15,000 forms. Associated with these potentials are pragmatic rules, with an average of 4 rules per item (around 600 rules), implemented as Lisp sublists of premises and conclusions. Semantic constraints are up to an average of 2 per item (around 300) implemented by the same method as pragmatic rules. Default rules do not directly belong to the lexicon, but are recorded in an interpretation module.

3.1 Descriptive Structure

The (semantic) **potential** of a word is its description in terms of associated concepts and their component features which specify these concepts with respect to the word. For instance, we think that in order to interpret sentence (1) (in introduction) there is a need to access a concept of '*hierarchy*' to be used to understand *father* process as hierarchically antecedent to the *son* process. Thus the concept of *hierarchy* is present to interpret father, and it is specialised (or featured) by the concept of *antecedence*, and the concept of *generation.*

Notation : let us call a major category such as *hierarchy* a **concept**, and let us call a featuring notion such as *antecedence*, which highlights the relationship between *hierarchy* and *father* in the context of sentence (1), a **feature**.

To every word w, we find in the lexicon :

$DS_{w1} = \{C_1, \ldots, C_n\}$ is the set of concepts associated with the word w.

$DS_{w2} = \{a_1, \ldots, a_t\}$ is the set of features describing the concepts recorded in DS_{w1}.

In order to create the proper correspondences between features and concepts we define a predicate B_w linking a particular feature to a concept in the context of the word w.

$B_w (C_i,a_j)$ = True if $C_i \in DS_{w1}$, $a_j \in DS_{w2}$ and a_j is a feature of C_i relatively to w.

$DS_w = DS_{w1} \cup DS_{w2}$. Any element of DS_w is noted x, when one doesn't want to distinguish between DS_{w1} and DS_{w2}.

$S_w = (DS_w, B_w)$ is the semantic potential of w.

Example

The following examples are extracted from our lexicon. They give an idea of the content of DS_{W1}, illustrating the fact that for the words we have recorded, only a few concepts (in the sense that we have defined above) are required to explain a word w.

$DS_{father1}$ = {hierarchy, affectivity, over-evaluation}, DS_{house1}= {building, dynasty, company}, DS_{sun1}={astronomy, social role, mythical figure}.

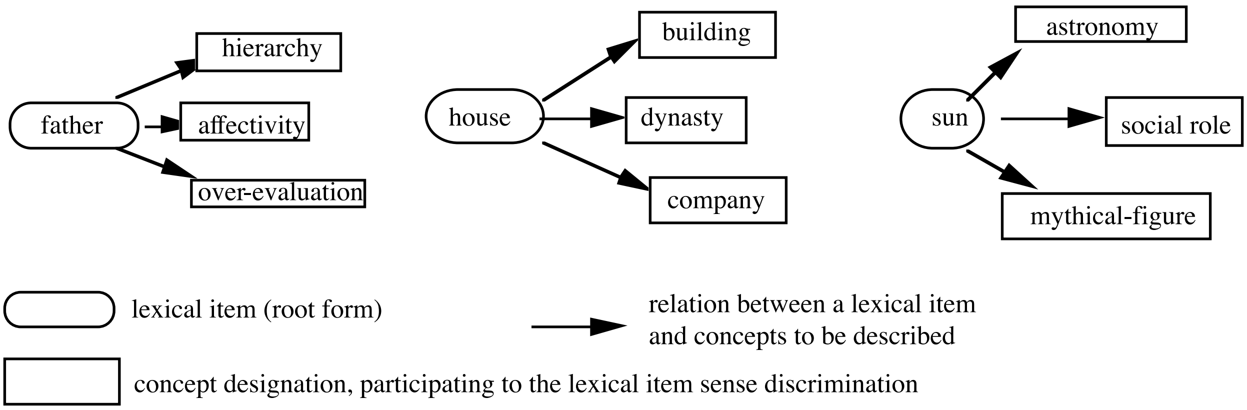

figure 1. Representing sense relationships

Note : at this level of description, lexical items are not linked with conceptual nodes bearing their own name. They are tied to concepts general enough to represent categories of notions, to be described from the 'point of view' of the lexical item usage, rather than already fine-grained concepts. This can explain that words such as *father, mother, parent, son, daughter, sister, brother, relative, wife, husband, child, family* are represented as tied with the same three generic concepts of *hierarchy, affectivity and over-evaluation* (that is sometimes represented in dictionaries as the idea of transcendance).

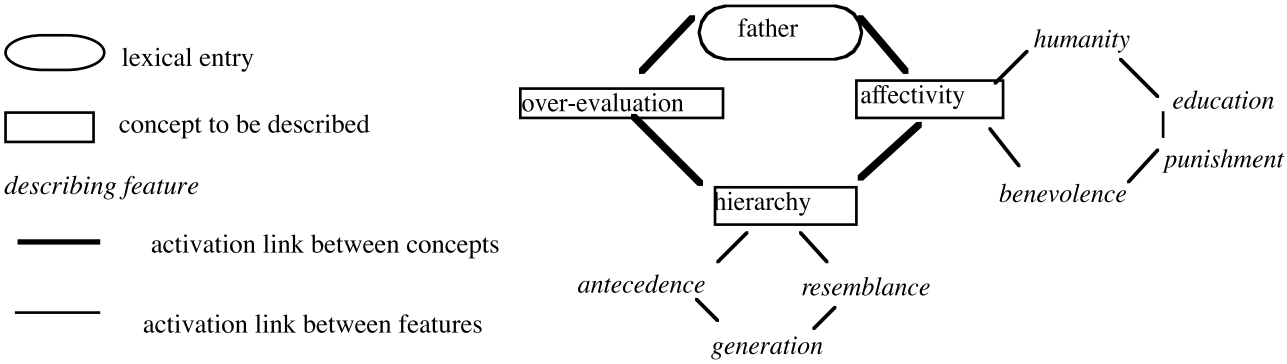

figure 2. A semantic potential

The representation of the complete structure, proposed in figure 2 for the word *father* as an example, indicates two things : first, how concepts are described in terms of relevant features (in the sense defined above), according to the lexical item which they are tied to. For instance, we will not find the same features of the concept *affectivity* associated with the word *father* as those associated with the word *child.* Second, the figure indicates links of a special nature between concepts: they are **activation links**. They mean that if a concept has been triggered by context then the other related concepts could be relevant to a lower extend, provided that no other information prevents it.

The example in figure 2 relies on the following information in the semantic potential of *father*, S_{father} :

$S_{father} = (DS_{father}, B_{father})$

$DS_{father} = (DS_{father1} \square DS_{father2})$

$DS_{father1}$ = {hierarchy, affectivity, over-evaluation}, represented by labels in boxes ;

$DS_{father2}$ = {antecedence, resemblance, generation, humanity, education, punishment, benevolence}, represented by labels in italics ;

B_{father} (hierarchy, antecedence), B_{father} (hierarchy, resemblance), B_{father} (hierarchy, generation), is represented by the 'cluster ' under the concept of hierarchy,

B_{father} (affectivity, humanity), B_{father} (affectivity, education), B_{father} (affectivity, punishment), B_{father} (affectivity, benevolence), is represented by the cluster under the concept of affectivity.

There is no feature for the concept of over-evaluation. There are two aspects to be explained :

- the links between concepts, represented in bold, and their circular representation : they act as activation links, for the flow of activity in context, which is to be detailed in § 3.2. The circular representation is justified in the same paragraph ;
- the links between features , represented by plain bars, and their circular representation ; they also act as local activation links, and are explained in next paragraph.

3.2 Inference and default rules

Inference rules deal with the activity rate of the semantic potential components: they define how relevant these components are in the interpretation context. In order to complete interpretation starting from a single piece of information, default rules are associated with the lexicon, and they are available for every described structure. Inference rules will act on the flow of activity in a semantic potential. Graphically their effect is on the links (see figure 2), both on the concept level and on the feature level. Activity is explained in § 3.2.1 and 3.2.2 where also the circular representation of figure 2 is justified, whereas rules are presented in § 3.2.3 and following.

3.2.1 Element activity rate

The **element activity rate** is a value that denotes to the extent to which the element participates in building the word sense in a given context. Let v be a value indicating the activity rate. v could point at a high activity (**salience**), an existing but hardly remarkable activity (the case for **validity** but not salience), absence of activity (**ignored**). We have also to deal with cases where salience is negated by context : we call this value **inhibition**. It helps to evaluate a negation marker, such as in :

(5) He is definitely ***not*** *a* ***good father****.*

Let V be the set of values to be assigned to features and concepts. V = { salient, inhibited, valid, ignored}. The values will be assigned by two main types of rules.

(i) Rules deriving activity from context : they assign a value to a node (concept or feature) as a result of the matching of their premises with information from parsing (see § 3.2.3) ;

(ii) rules 'spreading' activity : the values flow through the links (this is why we have to use a cluster-like structure with circular representation instead of hierarchical representation) on the feature level (by means of both default and semantic constraints, presented in § 3.2.4 and 3.2.5) and on the concept level (by means of simple default rules).

3.2.2 Rules of activity rate evaluation

Assumptions

Activity evaluation is launched by rules (mostly what we call pragmatic rules) and is completed when every node (except the lexical item, which is not considered as a node in our system) has been assigned a value. We function on the following assumptions :

(i) when a node (concept or feature) has received a value, then this value could spread, with a 'taxing' system, to every other node, thus being able to function with the least information possible

(ii) if a node receives different values (from different rules) then there is an order of priority to avoid inconsistency : specific values (such as salient, inhibited) are prefered to implicit values (such as valid)

(iii) the 'taxing' system will show: what node(s) are directly relevant to context (salient, inhibited), what nodes are implicit knowledge (valid), what nodes are totally irrelevant (ignored). If there is no direct relevance (nothing salient or inhibited) then implicit knowledge is a likely candidate for interpretation.

Model

In our system, rules are divided into three main groups :

Pragmatic rules, which translate the impact of contextual information (i.e. presence of other words in the sentence, morpho-syntactic features of the word, topic of the paragraph and focus of the sentence…) presented in §3.2.3.

Semantic constraints, which express regularities between features when they are activated : some features are linked by notions of opposition (such as between *good* and *bad*) and of implication (if *good* is salient, then *gentle* is also salient) , described in 3.2.4.

Default rules that 'spread' activity to the elements (concepts, features) of the potential that haven't been yet assigned a value by any of the semantic constraints or pragmatic rules, sketched in 3.2.5.

Both semantic constrants and default rules are rules 'spreading' activity and provide the components of our 'taxing' system.

3.2.3. Pragmatic rules

A **pragmatic rule** associated with a word acts as a launching rule for interpretation. Its premises are to be matched by context information and its conclusion is the association of a value to an element of the word potential.

Let w be a word and S_W= (DS_W, B_W) its potential. Let V be the set of values. Let p be a predicate defined as :

$$p : DS_w \; x \; V \text{ ------> } \{T, F\}$$

$$(x, v) \text{ ------->} p(x, v) = T$$

if v is assigned to x by a constraint (or a rule).

Let I_W be the domain of information known to be related to w (context information). An example of the content of I_W is provided hereafter. The set of pragmatic rules associated with w is :

$$P_{w1} = \{[\square j \; \square \; p(x, v)] \; \square j \; \square \; I_w, \; x \; \square \; DS_w, \; v \; \square \; V\}$$

Example

The following rule is relevant to the correct interpretation of the words '*mother*', '*father*', '*parent*'.

If the domain is technical	*general knowledge (in* I_{father}*)*
and w belongs to a noun phrase	*syntactic knowledge (in* I_{father}*)*
and w is directly prior to a noun	*syntactic knowledge (in* I_{father}*)*
then generation [in hierarchy] [in w] is salient	*value assignment to a feature*
and over-evaluation [in w] is ignored	*value assignment to a concept*
and affectivity [in w] is ignored	*value assignment to a concept*

This rule has been translated in Lisp using our formal representation. It has helped to understand the following uses of *mother* and *father* in the sentences we parsed:

(1) the **father** *process forks and generates a son process.*
(6) the **mother** *cell divided into two identical daughters.*

The results were the assignment of a relevant concept :

(1') the generator *process forks and generates a son process.*
(6') the generator *cell divided into two identical daughters.*

A similar rule was applied to the words *daughter* and *son*, emphasizing *descendance* (instead of *generation*).

3.2.4 Semantic constraints

Semantic constraints generally signal regular relationships between the components of the semantic potential of a word. We have collected corpus on common polysemous words : we have restricted our study to the words of parenthood, life (words such as life, breath, spirit, soul, death, essence and so on), and lodging (house, room, shelter, nest, etc.) in texts produced by 52 children (of the age 9 to 15), telling their view concerning an 'ideal world' (in [Prince 1986]). In this corpus, we have noticed that mainly two types of semantic constraints occur, and these are remarkable at the feature level only (i.e. in the set DS_{W2} of the semantic potential).

A feature A **implies** a feature B describing the same concept when B receives the same value as A in all contexts.

A feature A **opposes** a feature B describing the same concept when B receives the opposite value opp(v) of A, regardless of the context. We have defined an operator *opp* on V. The transformation table is the following :

v	Salient	Inhibited	Valid	Ignored
Opp(v)	Inhibited	Salient	Valid	Ignored

Let a_i, a_k be features of any concept C in the semantic potential of w. This can be translated by the following formula:

$$a_i,\ a_k \square DS_{w2}.\ \square C \square DS_{w1} / B_w(C,\ a_i)\ \&\ B_w(C,\ a_k).$$

Implication is defined as :

$$\square v_i \square V,\ a_i \square a_k \square [p(a_i,\ v_i) \square p(a_k,\ v_i)]$$

By convention, implication between features is defined using the same symbol □ as implication between terms in a formula.

Opposition is defined as :

$$\square v_i \square V,\ a_i / a_k \square [p(a_i,\ v_i) \square p(a_k,\ opp(v_i))]$$

By convention, we use the symbol / to represent opposition on the feature level.

The set of semantic constraints on a word w is given as :

$$P_{w2} = \{[[p(a_i,\ v_i) \square p(a_k,\ v'_i)] \square B_w(C,\ a_i) \square B_w(k,\ a_k)],\ C \square DS_{w1},\ a_i,\ a_k \square DS_{w2},\ v_i,\ v'_i \square V\}$$

Example

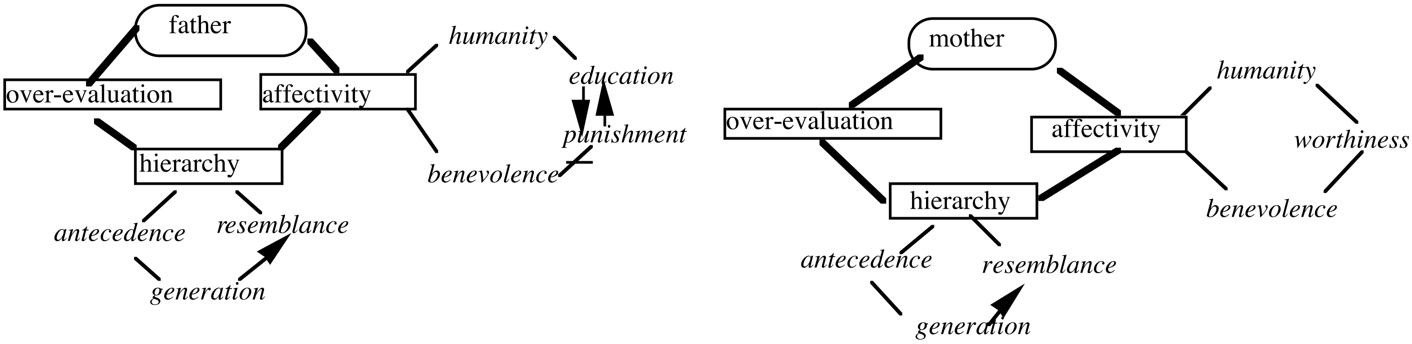

figure 3. Representing semantic constraints on two potentials of the lexicon

The semantic constraint of implication in the potentials of *father*, and *mother* between the feature *generation* and the feature *resemblance* in the *hierarchy* concept indicates that when *generation* is a likely feature in a context, then *resemblance* could also act as relevant information. It is represented in figure 3 by the arrow.

The semantic constraint of opposition in the potential of *father* is that the feature *benevolence* (describing the concept of *affectivity*) is opposed to the feature of *punishment* whereas the latter is implied by the feature of *education*..Please note that this is local to *father*. it is a cultural aspect of the father image in our society.

3.2.5 Default rules

Default rules deal with the spreading of activity. Five main rules have been formalised, but in this paragraph, we prefer to explain what they do in English, for the sake of clarity.

Conventions: we use *a* for feature, *c* for concept and *x* for any *c* or *a*. *v* is for value. *p* is the predicate that assigns a value *v* to a component *x*. *B* is the predicate that relates a concept to its describing features. We have removed the *w* in subscript because these rules are valid for all the words of the lexicon.

1. If x has received a value v and another rule tries to assign to it the value 'valid' then x keeps its preceeding value. Any other case is a problem of inconsistency between rules of the same type (i.e. two inconsistent pragmatic rules) and will make interpretation fail.

Example :

from the rule of page 5, generation receives the value 'salient'.

From the rule number 4 in this set, there will be an attempt to assign to it the value 'valid'. 'Salient' will prevail.

2. if a is a feature describing a concept c (we have B(c,a)), and a has received the value 'salient' or the value 'inhibited' then it passes on the value 'salient' to the concept c, if c has not yet received a value.

Example :

the rule in page 5 gives the value 'salient' to *generation.*

Thus the concept *hierarchy* receives the value 'salient'.

3. if a is a feature describing a concept c (we have B(c,a)), and a has received the value 'valid', or a concept c' has received a value different from 'ignored ' , and c' belongs to DS_{w1} *as c does, and c has not received a value , then c receives the value 'valid'.*

4. if a is a feature describing a concept c (we have B(c,a)), and c has received a value different from 'ignored' , then it passes on the value 'valid' to a (and this for every a such as B(c,a)).

Example :

The concept *hierarchy* , being 'salient' transmits the value 'valid' to its features.

5. if a is a feature describing a concept c (we have B(c,a), and c has received the value 'ignored', then it passes on the value 'ignored' to a (for every a).

Example :

In the rule of page 5, the concept *affectivity* , being 'ignored', transmits the value 'ignored' to its features.

4.Interpretation

4.1 Inputs and outputs: relationship with other modules

The input received from our interpretation module could be either of the following : an incompletely parsed sentence, with a morphological analysis, a syntactic analysis and a partial semantic representation (conceptual graph) provided by a deterministic parser ; a topic and focus discrimination; a reference resolution system. From this input, the system tries to match the premises of at least one pragmatic rule associated with the word were the parsing has stopped. If it succeeds, then it provides the following output : a probable candidate for the word sense if a particular feature has been found salient as the result of pragmatic rules application, a list of plausible candidates for the word sense, if a concept and/or a set of features have been found valid. Selectioning a candidate is made, by the interpretation module, as the result of calculating a word configuration in context.

4.2 Resolution algorithm

The **configuration** of a word is its interpretation in a given context. It is defined as a list of pairs where the first element belongs to the word potential and the second element is an activity rate.

Let w be the word to be interpreted. Let □ be the interpretation context (i.e. the sentence where w occurs, its topic, its focus, etc.). Let (DS_w, B_w) be the potential of w. The configuration □ of w in the context □ is defined as :

$$\square\ (w, \square) = \{\ (x,\ v), \square\ x\ \square\ DS_w,\ v\ \square\ V\}.$$

The configuration is obtained by applying the following algorithm:

```
For every node x in □ (w, □)) do /* while the configuration has not been completely calculated */
        begin
        while member-of ($\square_w$,□) do /* while a pragmatic rule premise matches an
        information of the interpretation context */
        /* apply pragmatic rules */
        apply $Pw_1 = \{[\square_j\ \square\ p(x, v)]\ \square_j\ \square\ \square_w, x\ \square\ A_w, v\ \square\ V\}$
```

```
end
 /* apply semantic constraints if they exist */
if $Pw_2$ <> □ then
apply $P_2 = \{[[p(a_i, v_i) \,\square\, p(a_k, v'_i)] \,\&\, B_w(c, a_i) \,\&\, B_w(c, a_k)],$
        $k \,\square\, A_{w1}, a_i, a_k \,\square\, A_{w2}, v_i, v'_i \,\square\, V\}$
/* apply default rules to finish configuration calculus */
apply P = { $R_i$, i = {1,2,3,4,5} & $R_i$ defined in §3.2.5} end
```

Example

The sentence :"the *mother* cell divided into two indentical daughters "

allowing the launching of the rule presented in page 5, has provided a configuration of the word *mother* (see semantic potential in figure 3). Two salient features have been found (*generation* and *resemblance*), one of which was valuated by a pragmatic rule, which has a priority upon the semantic constraint of implication (between *generation* and *resemblance*). Thus the most probable candidate is *generation* (in an agent case role, giving *generator*) and the other plausible candidates are, by order of plausibility, *resemblance* and *antecedence*. Both are considered as background information.

5.Conclusion

The lexicon we have presented in this paper has been used for word sense discrimination in unrestricted domain texts, when difficulties have arisen because of a 'non literal' use of some common words, mostly belonging to parenthood, breathing, stars and lodging terminology, corresponding to the themes of a collected a corpus. A further use of our model has been made by Ferrari (1992) to interpret lexical metaphors without creating a metaphoric link: the author has extended the lexical base to food and planets, using the same generic concepts, and has succeeded into interpreting unexpected metaphoric uses. We are presently studying the extension of our model implementation to other types of multiple senses, including cases of sheer homonymy (completely independent meanings associated to one word, such as *staff*/employees or *staff*/piece of wood), but mostly accounting for functional multiple senses such as *string*/of characters *string* /cord, *string*/music instrument. However, such a lexicon must not be evaluated independantly of other valuable knowledge sources such as morphological features or deep syntactic and semantic representations dealing with the sentence sense construction.

References

BRISCOE E., TAPPLETON A. 1991 "Some extensions as Lexical Rules ". *Proceedings of the IJCAI-Workshop on Computational Approaches to Non-Literal Language, Metaphor, Metonymy, Idiom, Speech Acts and Implicature*. University of Colorado at Boulder. Pp. 12-20.

CUMMING S. 1986 "Design of a master lexicon". USC/ Information Sciences Institute. Marina del Rey, California. RR 85-163.

EVENS, M.W., LITOWITZ, B.E., MARKOWITZ, J.A., SMITH, R.N., WERNER, O. 1983 Lexical Semantic Relations : A Comparative Survey. Linguistic Research, Inc, USA . Reed. 1986.

FERRARI S. 1992 "Etude des métaphores en traitement automatique du langage naturel." D.E.A dissertation (M.S.) University of Paris XI.

MARTIN J.H. 1988 "Representing Regularities in the Metaphoric Lexicon". *Proceedings of COLING-88*. Budapest. Hungary.

MARTIN J.H. 1990 A Computational Model of Metaphor Interpretation. Academic Press. New York.

McROY S. 1992 "Using Multiple Knowledge Sources for Word Sense Discrimination." Computational Linguistics, vol 18, n° 1.Pp.1-30.

PRINCE V. 1986 "Sémantique récurrente dans les langues naturelles et actualisation expérimentale". Doctoral Dissertation. University of Paris VII.

PRINCE V. , SABAH G. 1992 "Coping with Vague and Fuzzy Words: A Multi-Expert Natural Language System which Overcomes Ambiguities. *Proceedings of the Second Pacific Rim International Conference on Aritificial Intelligence (PRICAI-92).* Seoul, Korea.

PUSTEJOVSKY J. 1991 "The Generative Lexicon". Computational Linguistics vol. 17, n° 4. Pp. 409-442.

RIEGER, B.B. 1984 “ Semantic relevance and aspect dependency in a given subject domain”. *Proceedings of COLING84*. Pp. 298-301.
SABAH, G. 1990 “CARAMEL : A Computational Model of Natural Language Understanding using a Parallel Implementation”. *Proceedings ofECAI-90*. Stockholm. Pp. 563-565.
SMALL S.L., RIEGER C.J. 1982 "Parsing and comprehending with word experts (a theory and its realization)". In LEHNERT Wendy G. RINGLE Martin H. (Eds.) Strategies for Natural Language Processing. Laurence Erlbaum ass., Hillsdale NJ. Pp. 89-147.
SOWA J. 1984 Conceptual Structures : processing in mind and machine. Addison-Wesley. Reading,Massachussetts.
STALLARD D. 1987 "The logical Analysis of Lexical Ambiguity". *Proceedings of the 25th Annual Meeting of the ACL*, Stanford University, California. Pp. 179-185
WILENSKY R. 1986 "Some Problems and Proposals for Knowledge Representation". TR UCB/CSD 86/294. University of California. Berkeley.California.
WILKS Y.A. 1978 "Making Preferences More Active". Artificial Intelligence, n° 11. Pp. 197-223.