SEMANTICO-COGNITIVE REPRESENTATION OF MEANINGS
APPLICATION TO VERBAL POLYSEMY

Brahim Djioua
CAMS-LaLIC
96 boulevard Rapail 75006 Paris
djioua@msh-paris.fr

Abstract:
This communication describes the theoretical and practical tools that permit the automatic construction of a verbal lexicon in view of a didactic utilisation by linguists and an automatic treatment of written texts. Firstly we present the taken theoretical choices for the representation of the verbal polysemy in view of a practical utilisation; and secondly we present some elements of a semi-automatic tool (DISCC) that might help a semanticist in the construction of the semantic-cognitive representation associated with verb meanings.

Keywords: Verbal lexicon, polysemy, semantic-cognitive representation of meaning.

1 Methodological Preliminaries

Debates on the definition of word meaning is have been taking place for a long time in various complementary disciplines. Theoretical and practical studies have evolved in interaction with one other. Several authors [Victorri, 96], [Desclés et al, 98] recognise the importance of the lexicon in the semantic representations and note, in particular, that semantics is insufficiently used in natural language processing computer systems.

To handle a text written with a computer, artificial intelligence used semantic networks, the logic of predicates and the semantic representation with features. The utilisation of features semantics can be adapted to describe the nominal substantives but with difficulty for the verbal constructions because it cannot deal with verbal polysemy in charge. The semantic networks and the theory of prototypes is often put back in reason for problems on the categorisation and the typicality [Desclés et ali,91].

Syntax alone cannot reveal the semantics of lexicon. Verbal constructs with a same syntax can express different meanings. There are three examples in French::

(1) Pierre touches the cupboard [Pierre touche l'armoire]
(2) Paul receives a salary [Pierre touche son salaire]
(3) The garden adjoins the house [La jardin touche [à] la maison]

Every sentence has the same syntactic construction (4): NP1 VS NP2, but their semantic representations are different. The first sentence describes a simple contact between "Pierre" and the "cupboard"; the second expresses the entrance in possession of "a salary" by an agent entity "Paul". The third sentence implies the presence of a fictional observer that, while scanning the "garden", conceptualise more and more near places until to get in contact with the place "house".

HCP’99  Brest, France, September 1999
1.1 Processing polysemy

We can distinguish several methodological attitudes before the phenomenon of verbal polysemy. A first approach would be to consider that every senses of a verb is determined by the contextual environments ("effect of sense"), and therefore to refuse the oneness of "sense" for a verbal entity. The second attitude comes back to consider that, for every polysemic verb, there are as many distinct verbs and homonyms that of distinct senses [Atkins, 87]. Thus, the verb "to touch" would not have one senses but it exists several verbs "to touch" homonyms, each of this verbs are being identified by a semantic-syntactic diagram. This kind of approach can be applied for certain verbs homonymous like the French verb "voleer" : "voleer₁" (to fly with wings) and "voleer₂" (to take something without being allowed). Another attitude would be to consider that it exists a "sense core" that would be common to all uses of one same verb and every particular sense would be determined then by the context. In this case every sense is derivable of the "sense core", either by addition of semantic features, either by metaphorical operations. This attitude is often the one used by lexicographers for the construction of dictionaries that distinguish the different senses of one same verb by opposition of "concrete sense" to the "figurative senses ".

The approach adopted within the LaLIC team [Desclés,90] and [Abraham,95] is different of the presented previously attitudes. We suppose a comparable fundamental meaning to Picoche's "signifié de puissance" that would be to the source of the other senses. For Desclés, a polysemic verb has a "sense first" and its senses is analysed not from elementary linguistic predicates in the language in surface [Wierzbicka,92] but from more abstract elementary predicates, described in a cognitive representations level.

In this level of representation we uses some semantic and cognitive units (relaters and primitives) to construct some semantic and cognitive frames (SCF) associated to the verbal sense. [Desclés, 90] introduced a model, Applicative and Cognitive Grammar [A&CG], developed by the LaLIC team from the general works of the model of [Saumyan, 87] and of the cognitive extension of the model by [Desclés, 90]. The associate formalism to this model is a typed λ-calculus [Hindley-Seldin,86] and the Curry's combinatory logic with types [Curry-Feys, 58] with its combinators and its algebra. Indeed it has been shown by, among others, S.K.Shaumyan and by Desclés that combinators and types were excellent formal instruments of analysis linguistics, not only for the syntax but also for the verbal meaning representation [Abraham, 95]. This model articulates three levels of representations explicitly: a syntactic level, a predicative level and a semantic and cognitive level.

The first level is described by an extended categorical grammar [Biskri, Desclés 95]. The level two permits to formalise the predicative operations, of where of representations in term of predicates, arguments and role cases. The last level permits to represent meanings of predicates under shape of Sémantico-Cognitive Frames; these last are objects manipulated by our computer tool DISSC. The joint between levels is assured by a compilation process. Every level is described in a formal applicative language. The move from the middle to the lower level is a process of lexical integration.

1 For more details see [Desclés & alii,98]
SCF are semantic structures constructed with primitives. These primitives are defined by a basis semantic and cognitive types, static, kinematics and dynamic relaters and elementary operators as the topological operators on locations. The different static relaters are organised in a model of relaters more less specific of a location "archi-relateur" [Desclés, 87]. Relaters express different meanings (localisation, part-of, assignment, possession, identification,...) what one ordinarily designates in the semantic networks by the relation "is-a". These primitives have already been presented in several publications [Desclés, 85,90,98] [Abraham-Desclés,92]. SCF are described in a formal language with a specific grammar. Every SSC express a verbal meaning ; it is express either by a typed \(\lambda\)-expression, either by visual structures more accessible to the human users.

1.2 Semantic-Cognitive Frame Language

The conception of a design language supposes a definite grammar to basis of relaters and of primitive. The grammar of the SCFL defines a number of formal rules. The verification of the well formed applicative expressions is done by a functional typed system thus:

(i) Let some basis types \(\{I\text{ (individuals: Pierre, Paul, cupboard,...)}, C\text{ (collective)}, L\text{ (places: garden, house,...)}, M\text{ (massive), H\text{ (propositions), stat\text{ (static situations), cinem\text{ (kinematic situations), dynam\text{ (dynamic situations)}}}\) given

(ii) If \(\alpha\) and \(\beta\) are types then \(F_{\alpha\beta}\) is a functional type.

(iii) If X and Y are expressions of the type \(F_{\alpha\beta}\) respectively \(\alpha\) then the application \((XY)\) is of the type \(\beta\).

The static situations describe states. They are essentially described by the relation of location between a marked and a reference mark or by the relation of determination [Desclés,87]. Several static relations are defined between a reference mark and a marked. Relaters are typed; they are not compatible that with entities of a certain type. The static relaters, applied to typed entities, construct what one calls some static situations. Among the static relaters one finds the location "archi-relater" REP which specifies himself in different relaters : the differentiation, localisation and the assignment.

To explain the dynamism of a situation that is expressed of a static way, we suppose a fictional observer that, while scanning a global place, conceptualise some real or mental places witch are joined according to an any gradient.

Kinematics situations describe movements in a referential spatio-temporal or of changes of states assigned to an object. The kinematics primitives express a modification between two static situations. Among the primitive kinematics we use the primitive MOUV'T : the spatio-temporal movement of an entity passing a place to another (these places are not necessarily determined.)

The dynamic situations suppose an external constraint that gives back the possible kinematic modifications. The dynamic primitive express capacities of action from an individual entity in relation to a situation as FAIRE (make : capacity to do an action), CONTR (capacity to control an action) or TELEO (to aim a goal to reach). An operator
TRANS, got while composing the two dynamic operators CONTR and FAIRE, is used also to construct the dynamic situations expressing a semantic transitivity \( \text{TRANS} \equiv \text{W}_3 (\text{B}^2 \text{CONTR FAIRE}) \). Of another side the linguistic statement "Marie touche au but" is a predicative structure which can be reduced to a formal expression \( P_2 T_2 T_1 \); \( P_2 \) is a predicate with two arguments which expressed in the lexicon by the verb "toucher-au", \( T_1 \) the nominal term "le but" and \( T_2 \) the nominal term "Marie". By a compilation (a program of abstraction), we can reduce the symbolic expression \( P_2 T_2 T_1 \) to applicative expression \( \text{CONTR (FAIRE (MOUVTOR \quad SIT}_1[x,y] \quad SIT}_2[x,y]) \ x) \). The semantic meaning of the lexical unit "toucher-au" is an expression defined by the "lexical law": \( [P_2 \equiv B^2_2 \ W^3_3 \ B_2 \ \text{CONTR FAIRE MOUVTOR \quad SIT}_1 \quad SIT}_2 \ x] \).

To every polysemic verb is associated a network of SSC bound between them by relations of specification/generalisation, of abstraction or of domain change. For certain networks it exists a root that represents the common abstract significance to all the meanings of the same verbal item; this root is a cognitive "archetype" analogous to the "signifié de puissance" of [Picoche, 86].

2 DISSC: a tool to construct polysemic networks associated to a verb meanings

The construction, the manipulation and the consultation of the semantic-cognitive knowledge basis achieves itself in two big stages. The first task is to construct and to stock the SCF corresponding to the verbal items. A second stage consists in binding all its meanings by operations of fitting, be instanced or determination in order to construct networks between SCF translating the verbal polysemy. These two stages are processed very well separately either in parallel.

With this architecture (see figure 1), we recognise three levels of utilisation to the DISCC system.

• The administrator of the semantic-cognitive frame language, introduces and puts up to date relaters and the primitives of the SCF's language

• The semanticist can only create, modify and stock some SCF while using the only relaters and primitive definite by the administrator. The SCF, thus constructed, permit to construct networks of senses partners to polysemic verbs.

• The human user can consult the basis of SCF and networks. A system of text analysis that needs instances of SCF in a given context can either use it.

---

2 The complex combinators \( W_{(3)} \) and \( B^2 \) are respectively deferred and powers combinators. Deferred combinators are defined by a natural induction thus: \( X_{(0)} \equiv X \) and \( X_{(n)} \equiv B \ X_{(n-1)} \). Powers of combinators are defined by: \( X^1 \equiv X \) and \( X^n \equiv B \ X \ X^{n-1} \) [\( B \ x \ y \ z \geq x \ (y \ z) \)]. For more precision, see [Desclés,90].
The semanticist can, from a linguistic survey of meanings of a verb [Abraham,95] and [Desclés,98], to seize, to manipulate and to stock some SCF through the intermediary of the graphic publisher. The construction of SCF is made by a recursive way. The user chooses first a basis types, then static relaters to construct some static frames. At the end of the SSF construction, the DISSC system proceeds to a semantic-logical type verification. If the SSC constructs thus is coherent, the system generate its applicative representation automatically under shape of a lambda typed expression; otherwise a message of mistake is sent back. This verification consists in an inferential calculation of types on the applicative expression representing the SCF thus constructed. To a semanticist demand, an "integration program" permits to produce the lexical predicate with its typed arguments under the shape of a follow-up complex combinator of relaters, primitive and variables to be instanced by agents of the predicative relation.

While using the linguistic and cognitive analysis of the verb "to touch", we define all the SCF associated to the different meanings [Desclés, 98]. Then the polysemic network of the verb sense is constructed with abstraction / specification relations (see figure 2).

Research of information is done by a human user having access on the news visual more convenient than typed applicative expressions. Lexicon can be called also by a system as the automatic passage of a text on accidents reports toward a sequence of pictures [Battestelli-Valliez,97]. In some step of the semantic analysis of a text, it can need semantic value of a verb occurrence in order to "instanciate" a predicative relation.

This research in the lexicon can be done by a human user who want consult the lexical knowledge basis and see how are organised semantic categories of verbs. To title of example one can look for, for a given verb, and while taking account of the contextual information on the presence of a gifted agent of control in the predicative relation, all designs that are constructed according to certain constraints (with the primitive CONTR
The choice of a value on the "agentive" continuum (continuum d'agentivité) is based on a theory presented in [Desclés,94].

3 Conclusion

We presented here, in the cognitive setting of our research, a methodology of verbal lexicon construction. The polysemic verbal is used by an automatic treatment of the natural language. The DISSC tool is realised with the SCFL which use semantic and cognitive entities and typed primitives. DISSC is an interface programmed with Java™.

During the development of the system, our first goal was to develop an isomorphism between a graphic language and typed applicative grammar in order to have a reliable system. The system is under evaluation and a future work is to determinate the contextual rules witch permit the correspondence between a context and a verbal meaning.

References


Djoua, Brahim, Semantico-cognitive representation of meanings application to verbal polysemy; *HCP'99 (Human Centered Processes - 10th Mini EURO Conference)*; Brest, France 22-24 septembre 1999


[Desclés et ali,91] Desclés J-P and Kanellos I (1991), La notion de typicalité : une approche formelle, Sémantique et Cognition sous la direction de D. Dubois, CNRS


HCP’99 Brest, France, September 1999