The Translation of Basic Topological Prepositions from English into French

Nathalie Japkowicz

Technical Report CSRI-243
August 1990

Computer Systems Research Institute
University of Toronto
Toronto, Canada
M5S 1A4

The Computer Systems Research Institute (CSRI) is an interdisciplinary group formed to conduct research and development relevant to computer systems and their application. It is an Institute within the Faculty of Applied Science and Engineering, and the Faculty of Arts and Science, at the University of Toronto, and is supported in part by the Natural Sciences and Engineering Research Council of Canada.

The Translation of Basic Topological Prepositions from English into French

by

Nathalie Japkowicz

Department of Computer Science
University of Toronto
Toronto, Ontario, Canada
August 1990

A Thesis submitted in conformity with the requirements
for the degree of Master of Science

Abstract

Machine translation of locative prepositions is difficult, even between closely related languages such as English and French. We investigate translating the three prepositions 'in', 'on', and 'at' into the French prepositions 'dans', 'sur', and 'à'. Often, 'in' corresponds to 'dans' , 'on' to 'sur', and 'at' to 'à'. This correspondence, however, is not perfect: in a number of cases, the uses of these prepositions were observed to differ from one language to the other. These cases are not simply exceptional. Following recent work in cognitive science, we use the notion of *conceptualization* to account for this problem. A conceptualization (or metaphor) is a mental representation of an object or an idea. We believe that the differences in the uses of locative prepositions are caused by differences in the way objects are conceptualized in English and French.

We implemented a system that translates sentences involving one of the locative prepositions we studied from English to French, based on this idea. In addition, our system is able to detect ambiguities, as well as errors and abnormalities in input sentences.

Acknowledgements

I would like to thank my supervisor Janyce Wiebe for her inspiration and guidance. Her long hours and dedication were beyond the call of duty. Her suggestions and criticisms were always very helpful. I am also grateful to Graeme Hirst for his careful reading of this thesis and his many suggestions.

Thanks to all my friends and family for their help in discussing this work. Adam Deutsch, Michel Japkowicz, Julie Payette, Elana Ripsman, and Norrin Ripsman were particularly helpful. I especially would like to thank the members of my family, Pépé, Mémé, Papa (alias Zayde), Maman (alias Mamie), Florence, Maxime, Marc-Alain, and Joshua, for supporting me throughout my education and not hesitating to come to Toronto every time *they* felt they could help. Finally, thank you Norrin for *constantly* correcting my use of prepositions in English: these corrections were the starting point of this thesis.

I gratefully acknowledge the financial support of the Department of Computer Science, University of Toronto, and the Natural Sciences and Engineering Research Council of Canada.

Contents

List of Figures

Chapter 1

Introduction

1.1 Introduction

Machine translation of prepositions is difficult, even between closely related languages. The present study investigates the translation of three spatial prepositions, 'in' , 'on', and 'at', between two closely related languages: English and French. This subsection presents a series of illustrative examples followed by a discussion of the problem.

1.1.1 Examples

It seems at first glance that pairs of English and French spatial prepositions directly correspond to one another: 'in' and 'dans', 'on' and 'sur', and 'at' and 'à/au'[1]. In the following examples, they do indeed correspond to one another:

Example 1:

(1) Josh is *in* Dad's office.

(2) Josh est *dans* le bureau de Papa.

Example 2:

(3) Maxime was *on* this chair.

(4) Maxime était *sur* cette chaise.

Example 3:

(5) Mom is *at* her desk.

(6) Maman est *à* son bureau.

[1]The difference between 'à' and 'au' is simply morphological: 'au' is the contraction of 'à le' (preposition + masculine definite article).

However, these correspondences do not always hold. Sometimes, 'in', 'on', and 'at' are translated by 'dans', 'sur', or 'à', but the two prepositions are not corresponding ones. This is the case in the following sentences:

Example 4:

(7) Norrin is *in* the picture.

(8) Norrin est *sur* la photo.

Example 5:

(9) Marc-Alain is *in* the shade.

(10) Marc-Alain est *à* l'ombre.

Example 6:

(11) Florence is *on* the street.

(12) Florence est *dans* la rue.

Example 7:

(13) The poster is *on* the wall.

(14) Le poster est *sur le/au* mur.

Sometimes 'in', 'on', and 'at' are not even translated by 'dans', 'sur', or 'à'.[2] This is the case in the following sentences:

Example 7:

(15) Grandma and Grandpa are *in* France.

(16) Mémé et Pépé sont *en* France.

Example 8:

(17) The house *on* the lake ...

(18) La maison *au bord du* lac ...

Example 9:

(19) The tree *in* the road ...

(20) L'arbre *en travers du* chemin ...

[2]Some French prepositions or prepositional expressions other than 'à/au', 'sur' , and 'dans' can translate 'at', 'on', and 'in', in certain cases. These prepositions or prepositional expressions are: 'en' , 'au bord de/du', 'en bordure de/du', 'contre', 'à la jonction avec', and 'en travers de'. Examples of sentences using these prepositions or prepositional expressions will be shown in chapter 4.

1.1.2 Discussion

For some time, even to a certain extent into the eighties, the connection between language and spatial relations was considered straightforward — spatial relations in the physical world seem well defined and unambiguous. On this basis, geometrical and logical representations of natural language expressions of these relations were developed (Geometrical: Clark (1973), Hawkins (1983), and Bierwisch (1967); Logical: Cooper (1968), Leech (1969), Bennett (1968), and Miller and Johnson-Laird (1976)).[3] More recent studies (Lakoff (1982), Talmy (1983), Herskovits (1986), Vandeloise (1986), Mark (1990), and Grimaud (1988)), however, show that phenomena pertaining to fields other than logic and geometry are also involved in the linguistic descriptions of spatial relations. These studies suggest that natural language does not represent spatial relations as precisely and rigorously as strictly geometrical or logical representations do, and that it uses some kinds of information that do not appear in these kinds of representations.

One kind of enterprise that gives insight into the phenomena involved in the linguistic representation of spatial relations is the process of translation from one language to another. The present study will investigate translating the three prepositions 'in', 'on', and 'at' into the French prepositions 'dans', 'sur' , and 'à'.[4] As mentioned above, English and French are close to each other in their uses of these prepositions. Often, 'in' corresponds to 'dans' , 'on' to 'sur', and 'at' to 'à'. This correspondence, however, is not perfect: in a number of cases, the uses of these prepositions was observed to differ from one language to the other. These cases do not seem to be simply a few idiosyncratic cases. The problem is therefore interesting in that it motivates the search for extra information, other than geometric and logical, that could account for these differences. The present work proposes a systematic method for the translation of 'in' , 'on' , and 'at' from English to French, taking into account different types of information. Additional, even though closely related, goals of the present study are to detect ambiguity and semantically unacceptable input sentences.

The remainder of this chapter is organized as follows: section 1.2 introduces some definitions and some terminology. Section 1.3 describes the types of information that are determinant in the problem of translating 'in' , 'on', and 'at' from English to French. Section 1.4 explains how this information is organized in the system presented here. Section 1.5 describes the specific goals and strategies of our study. Section 1.6 presents an overview of the algorithm that will be used to perform the translations. Finally, section 1.7 gives the organization of the remainder of this thesis.

1.2 Definitions and terminology

Annette Herskovits (1986) calls the three prepositions 'in', 'on', and 'at', *basic topological prepositions* (Herskovits (1986), Chapter 9). We extended this term to the French prepo-

[3] See chapter 1 of Vandeloise (1986) for his criticisms of these works.

[4] These prepositions will be sometimes translated into other prepositions or prepositional expressions, as well. (See footnote no. 2.)

sitions 'dans', 'sur', and 'à'.[5] Basic topological prepositions have the property of being locative prepositions — they introduce an adverbial phrase of place and their meanings do not involve motion. In addition, they are distinguished from *projective prepositions* (Herskovits (1986), Chapter 10) — prepositions taking into account the point of view of the speaker, like in "The car is *in front of* the tree." — by the fact that they depend on just two parameters, the *located object* and the *reference object*, while the projective prepositions depend on these plus an additional parameter, the *point of observation* (Herskovits 1986). The located object is the object whose location is being specified in the sentence. The location of the located object is indicated in terms of the preposition and the reference object. For example, in

(21) The boy is at the supermarket.

'boy' is the located object and 'supermarket' is the reference object. In this sentence, the point of observation is not relevant because 'at' is a topological preposition. In sentences involving a projective preposition, the point of observation refers to the position of the speaker. Since this study is about topological prepositions only, we will not be concerned with points of observation.

Topological prepositions are defined in terms of syntax, semantics, and pragmatics.

Syntax

The main syntactic function of topological prepositions in English is the same as that of the other prepositions: linking two terms assuming different syntactic roles in the sentence. In the case of a topological preposition, the terms being linked together are either a verb and a noun phrase, where the noun phrase is the head[6] of an adverbial phrase of place, or two noun phrases, the first one playing any role in the sentence (subject, object, etc.) and the other one being the head of an adverbial phrase of place. We observed that the main syntactic role of topological prepositions is the same in French. It therefore does not account for the differences between English and French.

Syntax, however, might play *minor* roles in the differences in the use of topological prepositions between English and French. We found one such role in one domain only: the domain of geography. More specifically, the use of French topological prepositions obeys a syntactic rule in the case of the name of countries or regions. The French rule is that if the name of a country is feminine or if it is masculine and starts with a vowel, then the preposition 'en' must be used in the specification of location. Otherwise, 'à' is used (Gougenheim (1938) p.299-300). We illustrate this rule with a few examples.

Example 1:

(22) Marc-Alain is *in* Canada.

[5] French preposition 'en' might also be considered a basic topological preposition.

[6] We are using the term 'head' loosely. It is not related to the notion of 'head' in Chomsky's (1982) Government-Binding theory.

(23) Marc-Alain est *au* Canada.

Example 2:

(24) Grandpa is *in* Afghanistan.

(25) Pépé est *en* Afghanistan.

Example 3:

(26) Grandma is *in* France.

(27) Mémé est *en* France.

Examples 1, 2, and 3 illustrate the rule in the following way. 'Canada', 'France', and 'Afghanistan' are all names of countries. 'Canada' is masculine and starts with a consonant. It is therefore used with the preposition 'au'.[7] 'Afghanistan' is masculine, and starts with a vowel. It is therefore used with the preposition 'en' and no article. 'France' is feminine. It is therefore used with the preposition 'en' and no article.

Semantics

The semantic meanings of topological prepositions will be defined in terms of *ideal meanings*, following Herskovits: "The ideal meaning of a preposition is a geometrical idea, from which all uses of that preposition derive by means of various adaptations and shifts" [Herskovits (1986), p.39]. We will come back to this definition in chapter 2, where we will also identify the adaptations and shifts that Herskovits refers to. The ideal meanings of roughly equivalent basic topological prepositions of English and French were not observed to differ. However, the distribution of meanings was observed to differ in some peripheral cases.

We now illustrate what we mean by a different distribution of meanings. Russian and English offer us a clear example. In Russian, the prepositions 'in' and 'on' have single equivalent expressions but the preposition 'at' does not. In Russian,

'in' is expressed with 'vo' + locative case

'on' is expressed with 'na' + locative case

'At' is sometimes expressed as " 'vo' + locative case", and sometimes as " 'na' + locative case".[8] Probably because there is no equivalent of 'at' in Russian, the equivalent formulae for 'in' and 'on' assume meanings that are assumed by 'at' in English. Examples of such cases are the translation of the following English sentences or expressions into Russian:

[7] Recall that 'au' stands for 'à le'.

[8] 'At' might also be expressed with the locative case alone, or in other ways, but I do not know of any such use. This point is however not relevant to this discussion.

(28) He stood for a while at one place, then at another.

(29) In/at the agency, academy, swimming pool, cafe, museum, hairdresser's, university, etc.

and

(30) At the aerodrome, the post-office, the station, the factory, the farm, etc.

In sentence (28), the two uses of 'at' can be translated with either " 'vo' + locative" or " 'na' + locative". In the expressions of (29), 'in/at' is translated by " 'vo' + locative". In the expressions of (30), 'at' is translated by " 'na' + locative" (Wade 1983). This illustrates clearly the situation in which equivalent prepositions have different semantic distributions in two languages.

Such a phenomenon exists between English and French as well, although to a lesser extent than between English and Russian. There are probably fewer differences between English and French because French does have an equivalent to 'at'. From a general perspective, the main difference in the English and French distributions of the basic topological prepositions is that the French equivalent of 'at', 'à' , is in some cases used where 'in' or 'on' are used in English (but there are some other differences of this type as well). Additionally, French uses an extra preposition, 'en', in certain cases where 'in', 'on', or 'at' is used in English. Some of the examples of section 1.1.1 illustrate these facts. Section 4.3 of chapter 4 gives a more thorough analysis of the differences in the semantic distribution of the prepositions considered between English and French.

Pragmatics

A preposition can have different pragmatic implications in different contexts. Consider the following sentences:

(31) The fork is far from my hand.

and

(32) The post office is far from the school.

The same prepositional expression, 'far from', appears in (31) and (32), but has different pragmatic implications in each: in (31), the idea carried by the sentence is that the speaker cannot reach the fork without moving from his/her place, while in (32), the speaker might be saying that it is a long walk (or drive, depending on the context) from the post office to the school. Pragmatics is a relevant aspect of the use of topological prepositions, but will not be investigated in the present study. A reason is that our observations suggest that pragmatic implication is not an area in which corresponding topological prepositions differ in English and French. In addition, we analyze sentences in isolation, since this thesis is concerned with semantic meaning; thus, our analysis cannot contribute anything to the detection of pragmatically unacceptable utterances, which depends upon the larger context.

In summary, while the syntactic roles, semantic meanings, and pragmatic implications of prepositions are all important, we decided to limit our attention to their semantic meaning and to one syntactic rule. For the purpose of detecting unacceptable input sentences, this is an important restriction: the algorithm presented here considers sentences in isolation, and so cannot detect pragmatic anomalies. However, this does not appear to be a significant restriction for the purpose of translating topological prepositions from English into French, which is our main goal.

1.3 Factors in the use of locative prepositions

The previous subsection introduced some differences in the use of the three basic topological prepositions in English and French, which come from the prepositions themselves. These differences, however, are not major in the two languages considered. This section first briefly introduces what types of differences we do consider major, and then lists the factors involved in these major differences.

1.3.1 The major differences

The major differences, we claim following Grimaud (1988), come from differences in the *conceptualizations* of the objects involved in the relations. We will describe what a conceptualization is more specifically in the following section and chapter 2, but for now, we can understand a conceptualization to be the mental representation shared about an object by the members of a same culture (Lakoff & Johnson (1980))[9]. Following Grimaud (1988), we claim that English and French do not conceptualize objects in the same way and that this accounts for the major differences in the uses of the three basic topological prepositions between English and French. This point is the main point of this thesis. The main goal of section 1.3 is to draft a list of the kinds of information that we consider relevant to the way objects are conceptualized in modern North-American English and modern European and Quebec French.[10] The factors taken into consideration are the ones relevant to both our problem of translation and the detection of ambiguity and semantically ill-formed input sentences. The description of these factors follows.

1.3.2 The list of factors

The factors that we consider relevant to the uses of the basic topological prepositions in English and French belong to three different domains: geometrical, sociological, and natural. We first give an overview and then discuss each kind of information in turn. In the domain of geometry, we consider the following things to be relevant: the shapes of the objects involved in the relation, what parts these objects have, and the relative sizes of the two objects (for *normality checking*, which is described in section 2.3.1 of

[9]Note that this is what Lakoff & Johnson say about a *metaphor*. As will be clarified in section 2.2 of chapter 2, we use the terms *conceptualization* and *metaphor* interchangeably.

[10]Although my native French is Parisian, I believe that the analyses given in the present study are valid for other European French and Quebec French, as well.

chapter 2 and in section 5.2.2 of chapter 5). We consider one thing to be relevant in the domain of sociology: the *functions* of objects, that is, the ways that people interact with objects (Vandeloise 1986). Finally, we consider certain things to be relevant in the natural domain: some properties of objects. The algorithm presented in this thesis takes into account all of the above.

Shape

In some languages, the conceptualizations of objects are sensitive to the shape of the objects being considered. In English, for example, based on certain uses of the preposition 'at' , we claim that a reference object can be conceptualized as a point if it is a line (in addition, the located object must be fixed and the speaker must be on a trajectory intersecting the reference object[11]), as in

(33) The gas station is at the freeway.

This is not possible in French:

(34) * La station-service est à la route. (Herskovits (1986) p.86)

Parts

The parts composing the reference object are also a determining factor. In French, for example, we noticed that if the reference object has sides, then it is automatically conceptualized as a container, and 'dans' (the equivalent of 'in') must be used. This is not the case in English, which might conceptualize an object with sides differently. For example, French speakers say, in the case where a boy is riding a bus,

(35) Le garçon est *dans* le bus.

which is translated, in this context, as

(36) The boy is *on* the bus.

Social functions of objects

The social functions of objects are also determining factors in the use of prepositions. In English, for example,

(37) The boy is at the supermarket.

is fine, whereas,

(38) * The pigeon is at the supermarket.

[11]This will be made clearer in chapter 4.

is not, since a supermarket has the social function of being a place to shop, and a boy can have the social function of shopping but a pigeon cannot. This factor has the same effect in French. We considered this factor, however, in order to be able to address one of our subgoal, which is to detect sentences that are semantically unacceptable in a normal context.

Properties of the objects

Some properties of objects were also found to be important to the use of the basic topological prepositions in English and in French. The important properties are: the materiality of the objects, the material the objects are made of (note that these last two properties could be merged into a single one), their property to adhere to other objects, and the type of elements they are hermetic to. In some cases these properties are needed because they can account for differences between French and English in their use of prepositions, and in other cases, they are needed in order to accept or reject input sentences as semantically acceptable.

For example, the property of adherence is needed for translation purposes. In French, if there is an active adherence[12] relation between two objects, then the preposition used is 'sur' (Vandeloise (1986) pp. 200-202). Conversely, if the adherence relation is passive, then 'à/au' must be used. In the case of an intermediate adherence relation, both prepositions can be used. In English, the distinction concerning the adherence relation is not made, and 'on' is used in all three cases. We interpret the French distinction as follows: if there is active adherence, then French considers the two objects as being two agents independent from each other and occupying different positions in space. The preposition 'sur' is used in this case because, we believe, this preposition makes such a distinction. In the case of a passive adherence relation, French does not consider the two objects involved in the relation as active agents, and the distinction between them is not relevant to the relation. Their respective positions are not relevant either. The preposition 'à/au' is used because, we believe, this preposition carries this type of vagueness. In the case of an intermediate relation, either point of view can be taken and either preposition used.[13] Examples illustrating the phenomenon follow:

(39) The fly is *on* the ceiling.

can be translated only as

(40) La mouche est *sur* le plafond.

because a fly is an active agent in an adherence relation. However,

[12] ***Active, intermediate,*** and ***passive*** are Vandeloise's terms that qualify the type of adherence relation two objects are involved in. An ***active*** adherence relation is one in which the adherence is caused by one of the two objects (or both of them) involved in the relation. An ***intermediate*** adherence relation is one in which the adherence is caused by one of the two objects (or both of them) involved in the relation, plus a foreign object. A ***passive*** adherence relation is one in which neither object involved in the relation contribute to the adherence; instead, the adherence is caused entirely by a foreign object.

[13] This explanation is not a claim but is simply volunteered. It should be checked experimentally.

(41) The chandelier is *on* the ceiling.

can be translated only as

(42) Le lustre est *au* plafond.

because a chandelier and a ceiling are both passive. The adherence relation is caused by nails or screws: foreign agents. (Vandeloise (1986) pp. 200-201).

However, the property of hermeticity, for example, was needed for rejection or acceptance of the input sentence only, because the factor has the same effect in English and in French. This factor was used to reject sentences like

(43) The gas is in the glass.

on the basis that glasses are not hermetic to gases.[14]

Sizes

The relative sizes (and weights) of the located and reference objects are relevant for the purpose of detecting semantically unacceptable sentences. For example, we take into consideration that a sentence such as

(44) The bee is on the flower.

is fine, whereas a sentence like

(45) * The elephant is on the flower.

is not acceptable in a normal context[15], because an elephant is too heavy to be supported by a flower. However, English and French do not differ in this respect, so this factor is not relevant for the purpose of translation.

1.4 Organization of the information

All the information described in the previous section needs to be organized in a systematic way that allows translation between English and French. The organization we will choose is based on Lakoff & Johnson's theory of metaphors (1980), and Lakoff's organization of knowledge (1987). In this section, we give an overview of the organization of knowledge used in the system; more detail will be given in chapter 3.

We seek an organization of knowledge that would allow us to distinguish between the information that is needed for the type of translation we address — the translation of

[14]Note that this last remark is a simplification of reality, in which glasses can, in fact, contain gases that are heavier than air.

[15]Actually, the sentence is acceptable, if we consider that a crushed flower still qualifies as a flower. However, our default idea of a flower is of an uncrushed one.

topological prepositions from English into French — and information that is not specifically relevant. In addition, the knowledge should be organized in such a way that relevant pieces of knowledge are schematized, highlighted, and easily accessible. The representation that we find best suited to this purpose is based on Lakoff & Johnson's (1980) idea of metaphor. As will be explained in section 2.2 of chapter 2, the terms *metaphor* and *conceptualization* are used interchangeably throughout this study. A metaphor, as used by Lakoff & Johnson (1980), is a mental representation of an object, or an idea. This representation takes into consideration not only the "objective truth" about that object or idea, but also human biological perception and experience (sociological, historical, etc.). Metaphors are also a representation of humans' knowledge about objects and ideas, of the form we were looking for. Like Grimaud (1988), we think that if people conceptualize objects using the metaphorical principles suggested by Lakoff and Johnson, and that if, in particular, these conceptualizations depend on objective factors about objects, but also on cultural, historical, and sociological ones, then it makes sense that the speakers of different languages conceptualize objects in different ways, since they come from different cultures, have a different history, and possibly different sociological experiences. This difference of conceptualization could explain the differences in the use of the basic topological prepositions. This theory is supported by George Lakoff's (1987) theory on objects' multiple conceptualizations. Lakoff (1987) shows that many objects have multiple mental representations. The word 'window', for example, is one of these. a 'window' can be conceived of in two ways: as a hole in the wall, and as the glass-filled frames fitting into these holes. Two examples of the dual meaning of 'window' are:

(46) This room is too dark; we are going to have to cut a new window in that wall.

and

(47) They've just delivered our new windows.[16]

We now illustrate Grimaud's theory on translation between English and French with an example.

(48) Maxime is on the bus.

(49) Maxime est dans le bus.

In this example, we assume that the difference in use of the basic topological prepositions is due to the fact that in English, a bus is conceptualized as a surface, the *bottom platform* of the bus being highlighted, whereas in French, it is conceptualized as an empty volume, the *sides* of the bus being highlighted.[17]

Like Grimaud (1988), we assume throughout the present study that the difference in the conceptualization of objects is the main factor in the differences of translation of the three basic topological prepositions between English and French.

[16]This work of Lakoff (1987) p. 417-418 is discussed in Grimaud (1988), p.56.

[17]Note that sentence (48) is ambiguous in that Maxime can either be riding the bus, or be located on its roof. In this discussion we consider the first interpretation only, for the sake of clarity. However both cases are taken into consideration by the system, as will be shown in chapter 6.

1.5 The present study

The present study investigates sentences of the form:

⟨ Subject ⟩⟨ *be* ⟩⟨ Locative preposition ⟩⟨ Object ⟩

or noun phrases of the form:

⟨ Subject ⟩⟨ Locative preposition ⟩⟨ Object ⟩ [18]

The starting point of our study will be repeated: even though English and French are closely related languages, the translation of topological prepositions is not always direct. It seems at first glance that pairs of English and French topological prepositions directly correspond to one another: 'in' and 'dans', 'on' and 'sur', and 'at' and 'à/au'[19], but in fact, as the examples in section 1.1.1 illustrate, there are many differences in the use of the prepositions. Herskovits explains briefly that one source of difference between two languages that have the same central schema (such as English and French) is convention (Herskovits (1989), p.17). However, she does not extend the subject more. This work goes beyond her brief explanation.

The goal of the present study is to find some systematicity in the way the three basic topological prepositions are translated from English into French. As mentioned previously, this goal is accompanied by the subgoal of discriminating between acceptable and unacceptable uses of the prepositions in English. (The rules the algorithm uses to accomplish this would work as well for French input.)

As mentioned in section 1.4, we adopt the views of Michel Grimaud (1988), who says that the differences between English and French seem to be due to the differences of conceptualization of the nouns involved in the expressions, i.e., the located and the reference objects, in English and in French, rather than to a difference in meaning of the prepositions. Essentially, 'in', 'on', and 'at' correspond to 'dans', 'sur', and 'à' respectively (even though, there might be slight differences). What does not always correspond is the conceptualization of objects in the two languages. In a way similar to the way we explained the difference in sentences (48) and (49), Grimaud explains example 6 of section 1.1 (sentences (11) and (12)) as follows: in English, a 'street' is conceptualized as a roadway, whereas, in French, a 'street' ('rue') is conceptualized as a U-shaped container, including the buildings on both sides of the road (Grimaud (1988) p. 56). We will come back to this theory in chapters 2 and 3.

We take into consideration, however, the fact that every difference between English and French cannot be explained with differences of conceptualizations. One case where this is clear is the case of names of countries in French, as was discussed in section 1.3.2.[20]

[18] More complicated sentence structures would not bring any insight to the problem of translating topological prepositions from English to French.

[19] The difference between 'à' and 'au' is simply morphological: 'au' is the contraction of 'à le' (preposition + masculine definite article).

[20] Our system, however, does not make any distinction between cases that can be explained with differences of conceptualizations and others: the case of names of countries in French is treated by our

1.6 The algorithm

In this section, we first describe our view of translation, then give a brief step by step description of the algorithm (the algorithm will be described in greater depth in chapter 3), and, finally, we illustrate our algorithm with an example.

1.6.1 Our view of translation

Our method for translating sentences involving one of the three basic topological prepositions is based on the observation that a sentence has a central meaning which is language-independent, and a language-dependent mode of expression. The task of a translator is to express the language-independent central meaning of the source sentence in a language-dependent mode of expression of the target language. The modes of expression of the two languages may be the same or may be different. When they differ, the translator must choose a target-language mode of expression that underlines the same principal aspect of the situation as does the source-language mode of expression. The mode of expression in a language is the result of many historical, cultural, and sociological influences. The present work does not investigate why the mode of expression for a language-independent meaning may be different in English and French. It does, however, investigate the differences themselves, specifically those that are relevant to translating topological prepositions from English to French. We call the mode of expression of the language-independent central meaning of a sentence in a given language, the *conceptual level* of this sentence in that language. This conceptual level is made of the metaphorical entities talked about in section 1.4.

We recognize five different levels through which a source sentence passes until it is in its final form in the target language:

- The English language level
- The English conceptual level
- The language-independent level
- The French conceptual level
- The French language level

The decomposition into these different levels conforms to one of the standard machine translation methods: the *method of transfer*. The method of transfer works in the following way: it transforms source language into a source-language representation, which it then translates by a direct method into a target-language representation. This target-language representation is then used to generate target language. The transfer method is more powerful than the *direct method*, which translates a given source language into a given target language, using information specific to this given pair of languages, but

system, as if it could be explained with differences of conceptualizations. Chapter 7 discusses possible improvements concerning this problem.

not as powerful as the *interlingua method*, which performs translation by transforming source language into a language-independent form, and then translating this language-independent form into the target language. However, at the current stage of development of machine translation, the only operational systems use either the direct or the transfer method.[21] We illustrate our decomposition into different levels with Example 6 of section 1.1.

The English language level contains "Florence is on the street".

The English conceptual level contains "POINT SUPPORTED BY SURFACE".

The language-independent level contains "SPATIAL ENTITY SUPPORTED BY A PHYSICAL OBJECT WITH A PLANAR PART AND SIDES".

The French conceptual level contains "POINT INSIDE AN EMPTY VOLUME".

The French language level contains "Florence est dans la rue".

1.6.2 The algorithm

The system works roughly in the following steps:

Step 1: An English sentence is input to the system.

Step 2: The sentence is parsed, lexical ambiguity is detected, and the located and reference objects, the preposition, and the verb are identified.

Step 3: The French translations of each term except the preposition are found.

Step 4: All the English and French possible conceptualizations of the reference object are built. (We claim that the located object is always conceptualized as a point in both languages. We will come back to this claim in chapter 3.)

Step 5: One or more (in the case of conceptual ambiguity) English conceptualizations of the reference object are selected, according to the preposition used in the sentence.

Step 6: Candidate language-independent central meanings of the sentence are found, using the English conceptualization of the reference objects and checking if certain conditions in which such conceptualizations can be used are verified.

Step 7: The normality of the candidate language-independent central meaning is verified and the representations of the possible language-independent central meaning of the sentence are built.

Step 8: The matching of the representations built in step 7 with the corresponding conceptualizations of the reference object in French is done.

[21] For further information about the current stage of machine translation, see Slocum (1988).

Step 9: The French preposition to be used in the sentence is found using the French conceptualizations of the reference objects.

Step 10: The French sentence is generated.[22]

1.6.3 Example

We will now illustrate the operation of the algorithm on sentence (11).

Step 1: "Florence is on the street" is input to the system.

Step 2: The located object is 'Florence', the reference object is 'the street', the verb is 'to be', and the preposition is 'on'. Lexical ambiguity is detected at the level of the reference object.[23]

Step 3: The translation of 'Florence' is 'Florence', the translation of 'is' is 'est', and the translation of 'street' is 'rue'.

Step 4: In English, 'street' can be conceptualized as a SURFACE or as a FULL VOLUME[24]. In French, 'street' can be conceptualized as an EMPTY VOLUME or as a FULL VOLUME, given certain conditions.

Step 5: The English conceptualization of the reference object selected is SURFACE, because the preposition used in the sentence is 'on', and one of the ideal meanings of 'on' is "POINT SUPPORTED BY SURFACE"[25].

Step 6: There are no additional conditions for this candidate conceptualization to be instantiated.[26]

Step 7: The normality of the scene is verified, because a person can be supported by a street, and a person is a material object. The language-independent representation of the central meaning of the sentence is built. The property of the reference object that is highlighted in this representation is: "SPATIAL ENTITY SUPPORTED BY PHYSICAL OBJECT WITH A PLANAR PART AND SIDES".

Step 8: The French conceptualization of the reference object carrying the same central meaning is EMPTY VOLUME.

Step 9: The French preposition to be used is 'dans'.

Step 10: "Florence est dans la rue" is output.

---

[22]Note that this outline does not describe the entire algorithm since the handling of ambiguity and errors is not included. Section 3 briefly mentions the handling of these additional subgoals. Chapter 5 is entirely devoted to their description.

[23]There are in fact two translations of 'street' in French: 'rue' and 'chausée'. Our system takes this into consideration, but let us assume, for now, for simplicity, that the sentence is unambiguous, that is, that there is only one translation.

[24]The FULL VOLUME conceptualization carries the meaning that the located object is buried in the soil of the street.

[25]The ideal meanings of the three basic topological prepositions will be studied in chapter 3.

[26]See chapters 3 and 4.

1.7 The organization of the thesis

We end this chapter by presenting the organization of the remainder of the thesis. Chapter 2 overviews the earlier theories relevant for this study; chapter 3 presents in detail the algorithm that we designed; chapter 4 discusses the building of the conceptual information; chapter 5 discusses error and ambiguity handling; chapter 6 gives a series of examples of demonstrations of the system; and chapter 7 is the conclusion.

Chapter 2

Background literature

2.1 Introduction

This chapter gives an overview of the work in linguistics, psychology, and artificial intelligence that is relevant to this thesis. Our work is based mainly on Herskovits (1986), and Grimaud (1988). Before discussing their work in section 2.2.2, we will briefly describe in section 2.2.1 two important influences on these theories, Rosch's prototype theory of human categorization (Rosch 1977) and Lakoff & Johnson's theory of metaphors (Lakoff & Johnson 1980). Section 2.3 discusses some work that is peripherally relevant to this study, Vandeloise (1986), Mark, Svorou & Zubin (1987), Herskovits (1989), and Mark (1990). Section 2.4 explains how this study fits in with previous work.

2.2 The non-central theories

2.2.1 Human categorization: Rosch's prototype theory

We present Rosch's prototype theory (Rosch 1977) because one of the central ideas in Herskovits's work (described in the next section) is inspired by the prototype theory. As Herskovits says:

> There have been traditionally two ways to conceive of the meaning(s) of a lexical item:
>
> (a) The meaning of a word — assuming it has only one — consists of a list of conditions corresponding to its contribution to the truth-conditions of the sentences in which it occurs. ...
>
> (b) Another way to view lexical meaning involves the notion of prototype. [Herskovits (1986) p. 39]

Herskovits chose to define the spatial prepositions of English using the notion of prototype. We will now describe the prototype theory of Eleanor Rosch, and then we will explain how Herskovits was inspired by it. The prototype theory opposes theories that assume that the segmentation of the world is originally arbitrary. On the contrary, Rosch develops the idea that the world contains intrinsically different things. By this she means

that the world is naturally categorized. In her view, rather than imposing their organization onto the world, human beings categorize the world in a manner which conforms to the way the world is naturally organized. This idea is at the basis of her "Prototype Theory".

In the general attribute domains (such as the domains of shape and colour), Rosch argues that some perceptually salient, physiologically natural prototypes seem to determine categories. This means that some colors or some shapes are naturally more distinct than others and that colors and shapes can be categorized in terms of prototypes and amount of deviation from prototypes. She backs up this theory with a body of experiments performed on groups of people from different cultures and speaking different languages. The results show similarities in the way general attribute domains are categorized across cultures and languages, if the prototype theory is kept in mind.

She then tries to extend her theory to the domain of non-physiologically based categories of concrete objects. From her results, she argues that the contents of categories vary from culture to culture, but that the structure of the system of categorization is common to every culture and language. With a set of experiments (where each experiment is conducted within the same culture), she shows that this domain is structured analogously to the general attribute domain.

From these results, Rosch attempts to build a general theory of human categorization, in which objects are organized around prototypes, and are specified in terms of distances from these prototypes.

Herskovits uses this idea in her definition of the meaning of spatial prepositions: she gives each spatial preposition an *ideal meaning*, which is, as we said in the introduction, a geometrical idea from which all uses of that preposition derive by means of various adaptations and shifts (Herskovits (1986) p. 39) (described in section 2.3.1). These give rise to new relations that are in some way closely related to the ideal meaning. This "close relation" to the ideal meaning is similar to Rosch's notion of a "short distance" to a prototype.

2.2.2 Lakoff and Johnson's theory of metaphor

Metaphors, even though also not directly concerned with spatial relations, are discussed in this chapter because of the importance they take in our work. We base our study of translation of the three basic topological prepositions on the ideas of Michel Grimaud (his work will be described in section 2.3.), who relies extensively on Lakoff and Johnson's ideas. We will now describe Lakoff & Johnson's theory of metaphor, and will then discuss how Grimaud uses it.

George Lakoff and Mark Johnson built a theory of meaning which rejects theories based on "objective" or "absolute" truth. They show how human experience and understanding are more important to a theory of meaning than "objective" truth. The theories based

on "objective" truth consider metaphors as a matter of peripheral interest. Lakoff & Johnson (1980), on the contrary, give metaphors the central role. Metaphors are not formally defined by Lakoff & Johnson; however, we understand, from their claims that they extend the purely linguistic usual definition of a metaphor. They claim that metaphors are not a matter of language but that the human conceptual system is metaphorical in nature and that language only reflects this fact. They show, through linguistic examples, the mechanisms of metaphors and they attempt to build a theory of the human conceptual system.

Lakoff & Johnson extract the main mechanisms of metaphors. These mechanisms are:

- The systematicity of metaphorical concepts, which is the fact that if one concept exists in a culture, then linguistic expressions will systematically use this concept. For example, one metaphorical concept of the American culture is "TIME IS MONEY". A series of American expressions reflect this concept: "wasting time", "saving time", "giving time", etc.
- The hiding and highlighting of some aspects of a concept over others. This mechanism is illustrated by the following example. If the concept "ARGUMENT IS WAR" is considered (examples of this concept are "attacking arguments" and "winning arguments") then other aspects of arguments, such as the cooperative aspect in which someone helps someone understand his views, is totally hidden.
- The priority of certain concepts over others. This depends on the subculture from which the speaker emerges. For example, some subcultures give priority to the concept "BIGGER IS BETTER" over the concept "SAVING MONEY IS BETTER". Other subcultures do the opposite. The members of the first subcultures will have a tendency to buy big cars that are to be repaid in the future, whereas members of the second will have the tendency to buy small cars and save money.

Lakoff & Johnson then build a theory of the human conceptual system. They evaluate their theory by showing how, aside from explaining how language works, it accounts for problems that the older, more rigid theories of meaning could not explain and how it accounts for the evolution of language.

Grimaud (1988) uses Lakoff & Johnson's ideas (1980) to argue that if the objects involved in spatial relations are conceptualized in the way Lakoff & Johnson suggest, then coherence in the uses of the prepositions 'in', 'on', and 'at' within English, and their equivalents, 'dans', 'sur', and 'à', within French, can be found. Our study uses Lakoff & Johnson's theory on metaphors in the same way as Grimaud. In our work, following Lakoff & Johnson (1980) and Grimaud (1988), we consider that each object — in fact, category of objects[1] — has *mental representations*. A mental representation is the main

[1] For simplicity, we use the word 'object' for 'category of objects' throughout this thesis. For example, 'dog', in reality is a category of objects, and 'Spot' (who is a dog), an object. However, in this study, we call 'dog' an object and when we refer to entities, such as 'Spot', we think about them as individual members of a category.

general idea shared by an entire culture about an object. We also call the mental representations of an object its 'conceptualizations' (this term is borrowed from Herskovits (1986)). Mental representations, according to Lakoff & Johnson, are mainly metaphorical. In this study, we accept this view and, therefore, use the terms 'conceptualization' and 'metaphor' interchangeably.

We now present an example relevant to our study, which illustrates the use of this terminology. In the sentence

(50) Maxime is on the bus.[2]

In English, the two objects 'Maxime' and the 'bus' are conceptualized as a "point" and as a "surface", respectively. "Point" and "surface" are mental representations or conceptualizations of 'Maxime' and the 'bus', and, in the frame of our study, they are metaphors as well.

Let us now see how the metaphor "surface" in sentence (50) illustrates Lakoff & Johnson's mechanisms about metaphors. Some concepts of a 'bus' are hidden while others are highlighted: a bus has a bottom platform, which can be conceptualized as a horizontal surface, but it also has sides, a ceiling, wheels, etc., which are all hidden aspects when the metaphor "surface" is used for the 'bus' of sentence (50) (a similar analysis can be done with 'Maxime', who has a head, arms, legs, etc., but is conceptualized simply as a point). In addition, this example illustrates the priority of certain concepts over others: in English, the conceptualization of a bus as a "surface" has priority (assuming a usual situation) over its conceptualization as a container, for example. In French, however, the opposite is true and the corresponding French sentence is

(51) Maxime est dans le bus.

The use of 'dans' suggests that the bus is conceptualized as a container in French. The first of the mechanisms suggested by Lakoff & Johnson is not illustrated by our example.

2.3 Theories central to our study

Our work is based mainly on two theories: Herskovits's theory of spatial prepositions, and Grimaud's theory of the difference in the uses of spatial prepositions between English and French.

2.3.1 Herskovits's theory of spatial prepositions

Herskovits (1986) proposes a study of spatial prepositions deeply rooted in the works of Rosch (1977) and Lakoff & Johnson (1980). Like them, Herskovits rejects the classical

[2] As noted in section 1.4 of chapter 1 when the same example was discussed, the sentence is ambiguous: Maxime could be riding the bus or could be located on the roof of the bus. We concentrate, for the sake of this demonstration, on the first meaning. The ambiguity is taken into consideration by our system, as will be discussed in chapter 5.

categorization of the world based on the "objectivist" view of meaning, which she claims is not adequate to describe the spatial domain.

She creates a theory for the treatment of spatial prepositions which relies on ideas similar to those of Rosch: the idea of prototypes, or in the case of spatial prepositions, the idea of *ideal meanings* . The central goal of Herskovits's study is to deal with the *decoding/encoding* problem. The decoding problem consists of finding the spatial representation described in a natural language sentence and the encoding problem is the opposite problem of describing a spatial situation in a natural language sentence.

Our problem of translating the three basic topological prepositions from English to French is, similar to the decoding and encoding problems. Our problem is divided into two phases. The first one can be compared to the decoding problem; it must derive a non-linguistic representation of the meaning of an input sentence. The second phase can be compared to the encoding problem; it is to derive a natural language sentence that expresses a non-linguistic representation of a situation. The differences between our system and solving the decoding and encoding problems are the following. First, obviously, our system contains an extra module that performs transformation from one language to the other in order to translate the sentence, and second, our system derives a less detailed representation of the scene than the decoding and encoding problems would (if they were implemented). We do not have detailed representations of the scenes because such detail is not needed for the purpose of translation between English and French. The reasons, we argue, are that the differences are mainly conceptual in nature, and that the two languages express spatial relations to a comparable level of precision.

Herskovits's analysis of the locative expressions in English is based on the observation that spatial prepositions are complex elements of language and simple geometric relations are not sufficient to represent them. She proposes a new theory based on the observation that a spatial preposition has primarily a geometrical meaning, or the *ideal meaning* of the preposition, and that this ideal meaning can be stretched in various ways, leading to deviated acceptable meanings. She recognizes three main bodies of deviation from the prototypical meaning: the *sense shifts*, which are conventionally allowed deviations; the *tolerance phenomena*, in which a prototypical meaning or the result of a "sense-shifted" prototypical meaning are allowed to be considered "almost true"; and the various transfers to geometric description in which *synecdoche* (using a more inclusive term than is required) and *geometric conceptualization* (conceptualizing objects as geometrical elements) play a considerable role.

As discussed below, we do not recognize the sense-shift body of deviation. Tolerance phenomena do not play a role between English and French: our assumption suggests that tolerance phenomena are the same in both languages with regard to spatial prepositions. The various transfers to geometric description is the kind of deviation we concentrate on and, in particular, the geometrical conceptualizations.

We now give some of Herskovits's terminology, and indicate which of the objects she

described are actually implemented in our system.

The main notions of Herskovits's work are the *ideal meaning*, which is the prototypical meaning of a preposition, the *use types*, which are deviated meanings from the ideal meaning through sense shifts, and the *normal situation types*, which can be defined in terms of the use types as follows: a use type can be seen as a set containing an infinity of situations or scenes known as the *situation types*. (Note that Herskovits believes that all situations have to be described by a use-type.)[3] Among this infinity of situation types, some describe ordinary scenes, such as:

(52) The poster is on the wall.

and others, out-of-the-ordinary scenes, such as

(53) The woman is on the wall.

The situation types describing ordinary scenes are called the *normal situation types*.

We will use the notions of ideal meaning and normal situation types in our study, although we redefine these terms slightly. We will come back to this point when we describe the algorithm in chapter 3.

As mentioned above, we disagree with Herskovits's point concerning the sense-shift body of deviation. Herskovits argues that use of the topological prepositions is conventional to a great extent and that topological prepositions have an ideal meaning, together with a list of conventionally derived meanings attached to them, the use-types (Herskovits 1986; 1989). We do not agree with this view. In our view, the prepositions have one or several ideal meanings, but no use-types attached to them. We recognize that convention plays an important role in the use of topological prepositions, but not at the level of the prepositions. Convention, in our theory, intervenes in the way the objects are conceptualized. For example, Herskovits quotes the following difference between English and French (Herskovits (1989) p.17)

(54) he sat in/on a tall chair.

(55) il s'assit *dans/sur une chaise.

She argues that such a difference argues in favor of her idea that topological prepositions have a list of meanings, and that the French list differs in certain places with the English one. We disagree with this explanation, and argue that the fact that 'dans' cannot be used in the French sentence is due to the fact that a 'tall chair' and a 'chaise' are not always conceptualized in the same way in the two languages: they both can be conceptualized as a flat surface *on* which people can sit, but only the 'tall chair' can be conceptualized as a container, *in* which people can sit. French actually has a special word for the type of chair *in* which people can sit: 'fauteuil' ('armchair' is the closest

[3] We believe otherwise, as will be discussed below.

translation). The same argument holds for other objects where there is a difference of conceptualization between English and French, such as buses or shade.

As we said earlier, however, there are some uses that are not motivated but can be considered as conventional, as in French with proper names of countries or regions. As we saw earlier, however, these uses are motivated by syntactic rules, morphological rules, or historical reason, but not by purely cognitive principles.

2.3.2 Grimaud's theory on the cross-linguistic differences in the use of spatial prepositions

Grimaud (1988) is a cross-linguistic study of toponyms, prepositions, and cognitive maps whose goal is to point to differences in spatial relations among languages, in order to provide insight into the organization of languages in the spatial domain. The comparison is made on three languages or dialects: French, American English, and British English.[4] Throughout the paper, Grimaud conducts analyses based on the ideas of Lakoff & Johnson and Herskovits described above. The paper discusses the distinctions in the uses of 'in' and 'on' in the three languages or dialects, when applied to streets, street names, or other geographical entities. Grimaud suggests that there is coherence in these uses when the objects of the spatial relation are conceptualized, as Lakoff & Johnson describes it. He also distinguishes the uses of 'in' and 'at' when used with names of buildings and countries. Here again, he suggests coherence, but this time with respect to the functional description of the objects: the differences, he claims, come from the distinction of social function versus geographical location.

We first look at the 'on' versus 'in' distinction. In discussing 'on' versus 'in' in French, American, and English, Grimaud concentrates in particular on streets, street names, and on the metaphorical use of 'on', as in "The house is on the lake". From the observation that English people say "...on the street", and French people say "...in the street", Grimaud suggests the explanation already given in our introduction, that the English language focuses on the roadway meaning of streets, whereas the French language conceptualizes streets as a kind of U-shaped container, including the buildings on both sides of the roadway. In order to show some foundation for this explanation, Grimaud gives an argument in three parts: he first shows that multiple mental conceptualizations are not unusual, by quoting some of Lakoff's examples (Lakoff 1987); in a second argument, he shows, with examples, that English and French use spatial prepositions in a closely related way; and in a third argument, he shows, with examples, that in the right context, English uses 'in' instead of 'on', and French 'sur' instead of 'dans'. The idea of Grimaud is summarized by Bennett (Bennett (1975) pp. 66-67), who says that the physical characteristics of objects play an important role but are not sufficient in determining entirely the speaker's usage: the important point is the way objects are thought of on particular occasions. This idea will be used very much in our work, where the same object will

[4] In our study, we do not consider British English, but we do consider North American English, and French (European French and Quebec French, which we believe use topological prepositions in the same way, except in the cases discussed in section 4.2.2.3 and 4.2.2.4 of chapter 4).

be given several possible conceptualizations, given different contexts (see chapter 4). A fourth argument defending Grimaud's theory would be that of systematicity: if many uses of topological prepositions could be explained using Grimaud's theory, then this theory would be reinforced. The present work applies Grimaud's theory systematically, on many different types of sentences, in an attempt to reinforce this theory.

With regards to metaphorical uses of the preposition 'on', such as in "The garden is on the lake", Grimaud observes differences between English and French. He gives sufficient but not necessary reasons for these differences: in English, 'on' can be used when the reference object can be conceived of as a line, and when the located object is on either side of the reference object, and can be named in relation to this line. The preposition 'on' can also be used if the located object straddles the reference object. In French, it is more difficult to conceive of objects as lines and to relate items on either side of the reference object to that line; therefore, the use of 'sur' is less extended than in English in this direction. The only use of 'sur' that is allowed, in this framework, is the one where the located object straddles the reference object. The reasons why 'sur' cannot be extended as in English were not found by Grimaud, who suggests research in historical evidence that would motivate this fact.

We use the work of Grimaud for the 'on' versus 'in' distinction in our theory and extend his theory to objects other than purely geographical ones. For example, we deal with vehicles, such as buses and planes, as well (Grimaud actually mentions them too, but does not compare them with French).

In his discussion of 'in' versus 'at', Grimaud notices that in the cases of buildings, the use of the two prepositions is very close in English and in French, but in the case of geographical entities such as countries, cities, and regions, they are different. In order to explain the use of 'in' versus 'at' in the cases of buildings both in English and French, Grimaud introduces the notion of the functional meaning of spatial prepositions. Grimaud shows, with an example, the difference between 'at', 'in', and 'in the'. The example he uses is 'prison'. Here are Grimaud's words:

> The first construction ("he is at the prison," "il est à la prison") indicates mere location specified in answer to a question like "Where is he?" The prison is a point on a map. The second example ("he's in the prison," "il est dans la prison") indicates focus on the prison as an individual building conceptualized as a space *within which* one is moving. The third construction ("he's in prison," "il est en prison") indicates that the man is a prisoner.[Grimaud (1988) p. 65]

Moreover, an extra meaning of 'at' and 'à' that does not apply to 'prison' is the aspatial functional use of 'at', as in "Where you at the theater yesterday?" ('Etais-tu au theatre, hier?'); the speaker actually wants to know if his interlocutor saw a play, the day before. The use of 'in' in the sentence above would emphasize the spatial aspect instead.

In our study, we also distinguish the uses of 'in' and 'at', using Grimaud's remark,

by specifying the social function of some buildings or areas (such as a swimming pool) in our knowledge base. As well, we include specific information about the type of activity each located object can perform: a baby cannot study at the university, but a young man can, for example. In the case of geographical entities, the uses are not as clearly dependent on functionality principles; rather, the use of prepositions in French depends mainly on their history and morphological constructs. Some uses with geographical entities, however, can also be functional; in "The Marriott [hotel] at Copley Place", for example, 'at' is used because Copley Place is luxurious. In our study, the use of prepositions with geographical entities, such as cities, countries, etc, in French, will be frozen.

2.4 Other works

This section discusses other works that are relevant to our study, but less so than the works previously described. These works are relevant in one of three ways: either they provide us with general directions, they provide us with case studies of the uses of the three topological prepositions, or they provide us with some methodological tools useful for our theory. Vandeloise (1986) and Herskovits (1989) give us some extra general directions for our theory, Herskovits (1986)[5], Vandeloise (1986), and Mark (1990) provide us with case studies of the use of the three basic spatial prepositions in English and French, and Mark et al. (1987) provides us with some methodology.

2.4.1 The general directions

All the studies mentioned in this section argue that the treatment of space ought to take into consideration some of the cognitive principles inherent in human beings. This view is important to our study (the cognitive principles in question are the function of the objects and the conceptualizations of objects by human beings, that were described in chapter 1 section 1.3). Vandeloise (1986) and Herskovits (1986) both reinforce some of these principles.

Vandeloise (1986) advocates the view that in order for the uses of spatial prepositions to be well understood, the objects involved in the relations have to be considered from a functional point of view — that is, in the ways humans interact with them. Vandeloise wrote a linguistics study and decomposed the problem to a greater extent than we did; as well, he dealt with a greater number of prepositions than we did. However, his idea was kept in mind in the development of our study and we also took into consideration human interaction with objects.

Herskovits, in her recent article (Herskovits 1989), summarizes the principles that the use of prepositions abide by. These principles are also cognitive ones. These principles are:

- The structural description of shape, which she defines as the relations between the parts of the objects involved, the properties of these parts, and of the whole

[5]In addition to her theory already summarized in section 2.3.1.

objects. The types of properties and relations investigated are: symmetry, parallelism, elongation, straightness, curvilinearity, horizontality, etc.

- Perceptual grouping, which is the ability of the human perceptual system to group objects in the visual field, given their proximity, similarity in size, colour, etc., making them into one entity.
- The representations used in motion planning, which is the human ability to idealize certain objects to geometric entities (points, lines) when a trajectory is inherent to the relation described (Herskovits assumes this principle, even though she admits that she has no psychological backing for it).
- The processes of selective attention, which is the human ability to select certain parts of the objects over others, and give them an active part in the expression of the relationship, while hiding the other aspects of the objects.
- Location representation and frames of reference, which say that in order to understand projective prepositions — prepositions which link together three entities: the located object, the reference object and the observer — cognitive knowledge of the objects has to be taken into consideration.

The four first points are actually being used by our theory. The last point would be useful if the projective prepositions were investigated.

2.4.2 The case studies

In addition to the ideas discussed in section 2.3.1 and 2.4.1, the work of Herskovits (1986) also provides us with case studies of the use of the three basic topological prepositions in English. Her case studies are well detailed and give us some of the motivations for why a specific preposition is used in a certain case. Our use of prepositions is mainly based on Herskovits's case studies.

In addition to the general direction discussed in section 2.4.1, Vandeloise (1986) provides us with case studies of the uses of two of the three basic topological prepositions in French: 'sur' and 'dans'. Some of the motivations for the uses are also given. The case studies are, however, not as detailed as Herskovits's, since Vandeloise tries to give a "general idea", whereas Herskovits takes a more detailed approach.

The work of Mark (1990) reviews one of Herskovits's cases and makes some modification to it. His study deals with six different languages, and in particular English and French.

We use these three works in our study.

2.4.3 Methodological tools

The paper of David Mark, Soteria Svorou & David Zubin (1987) provides us with methodological tools concerning the division of objects and their surroundings into regions. Since

the authors study more prepositions than we do, the extent to which they divide these objects is greater than we need; however, we borrow some of their ideas. Their decompositions are as follows. Each object has

- An interior
- A surface
- An exterior area adjacent to the object.

The size of the exterior area is a problem because it is not bounded by the object itself. Instead, the exterior region has a scale determined by

- The size of the object itself
- The distance to, and/or size of, other relevant objects
- A functionally-determined "use-space" around the object

As well, the exterior region of an object is partitioned in two ways:

- A one-dimensional vertical space: a subregion above and below the object and a two-dimensional horizontal space around the object
- A four-way division of the horizontal space corresponding to the front, back, right, and left.

From their work, we borrow the subdivisions into the interior, the surface, and the exterior regions adjacent to the object, as well as the size of the object itself, the distance to, and/or size of, other relevant objects, and the functionally-determined "use-space" around the object. We do not need the other two features, which might be useful for computer representations of the relations (which are not very detailed in our case, since as mentioned in section 2.3.1 detail is not needed for translation between English and French) and for the translation of the projective prepositions.

2.5 Our work in perspective

This section places our work in perspective with previous works in the disciplines of linguistics, psychology, and artificial intelligence.

2.5.1 Our work in perspective with earlier linguistic and psychological work

As was mentioned in section 1.1, our study follows some linguistic and psychological studies concerning language in connection with spatial relations. Two main trends can be identified in these works: the works defining geometrical representations of this relation (Clark 1973, Bierwisch 1967), and the works defining logical ones (Bennett 1968, Miller and Johnson-Laird 1976). We use some of the principles described by the geometrical and logical representations, together with the cognitive principles that were discussed earlier.

We briefly talk about a few points of Clark (1973) and Bierwisch (1967). Clark defines 'at', 'on', and 'in' as uni-, bi-, and tri-dimensional, respectively. That is, 'at' determines relations between points and/or lines, 'on', relations between points, lines, and/or surfaces, and 'in', relations between points, lines, surfaces, and/or volumes. Bierwisch talks about the priority of certain dimensions[6] of objects over others, in certain situations. That is, in certain situations, one of the dimensions might be more relevant than another in the relation.

These works have been criticized for their inadequacy in describing spatial relations in language (see Vandeloise (1986)). However, we use ideas similar to these of Clark and Bierwisch in our algorithm. We define 'at', 'on', and 'in' in terms of uni-, bi- and tri-dimensional relations, but we add some cognitive principles to these definitions. That is, we define the prepositions as uni-, bi-, and tri-dimensional relations between *conceptualizations* of objects, rather than between geometrical description of these objects themselves. Similarly, we use the idea of the priority of certain dimensions over others, like Bierwisch does, but we integrate this principle with mechanisms for deriving conceptualizations.

Logical representations of spatial relations were also attempted in order to describe spatial relations in natural languages. For example, Bennett (1968) and Miller & Johnson-Laird (1976) give the following definitions of 'in':

Bennett: in y= locative (interior y)

Miller & Johnson-Laird: In (x,y)= referent x is in a relatum y if part(x,z) & Incl (z,y)

These descriptions have been criticized as not powerful enough to describe all the spatial relations in language (see Vandeloise (1986)). We, however, use definitions similar to the ones above for the ideal meanings of 'in', 'on', and 'at', in our study. The difference between our use of logic and the use of logic in the works previously described is that here again, we added a cognitive element by saying that the ideal meanings of the prepositions apply to the conceptualizations of the objects rather than to the objects themselves.

The reason why conceptual descriptions are more powerful than pure geometrical and logical ones is that conceptual descriptions carry in them selected knowledge belonging to several different domains, whereas geometrical and logical descriptions carry information belonging to a single domain.

2.5.2 Our work in perspective with works in artificial intelligence

Two types of work in artificial intelligence deal with spatial relations.

[6]The term 'dimension', here, refers to the width, the length, and the thickness of objects.

The first type of work deals with spatial cognition alone (not related to language). The work on cognitive maps (Kuipers 1978), for example, belongs to this category. This work tries to understand human beings' orientation in a large-scale environment. Since this work deals with movement, points of reference, and large scale environment, it is not closely related to our study. We will therefore not discuss it here.

The second type of work is that dealing with the problem of specifying spatial positions using natural language. For example, the work of Herskovits (1986, 1989), Mark (1990), and Retz-Schmidt (1988) belong to this type. The work of Herskovits and Mark relates to this study and was already described in sections 2.3 and 2.4. Retz-Schmidt (1988) is interesting in that it actually describes an implemented system, CITYTOUR[7]. CITYTOUR deals with the relationship between language and vision. However, CITYTOUR deals with large-scale environments, and not with detailed objects. In the system, objects are simply given a general shape and are not given much attention in other respects. In addition, the system focuses on projective prepositions, and therefore, does not bear much relevance to our study.

[7]Herskovits did not implement her study, and Mark's implementation is in progress; therefore, aside from CITYTOUR, I have not come across any system dealing with spatial relations in natural language.

Chapter 3

An overview of the algorithm

3.1 Introduction

This chapter presents an overview of the algorithm for translation of topological prepositions. In Section 3.2, we describe the tools used in the implementation of the algorithm. In section 3.3, we outline the basic modules of the algorithm. Section 3.4 describes the main entities that the algorithm operates on. Finally, in section 3.5, we describe the modules of the algorithm in greater detail.

3.2 Tools used in the implementation of the algorithm

The algorithm is implemented using the Semantic Network Processing System (SNePS) (Shapiro 1979; Shapiro & Rappaport 1987; Shapiro 1989). SNePS is a system which can build, use, and retrieve from, propositional semantic networks. A semantic network consists of labeled nodes, which represent concepts, and labeled directed arcs, which represent binary relations between concepts. An arc labeled R going from node n to node m represents that the concept represented by n bears the relation represented by R to the concept represented by m (Shapiro 1979; Shapiro & Rappaport 1987). SNePS is called a *propositional semantic network*. This means that every proposition represented in the network is represented by a node, not by an arc. In addition, SNePS comes with some packages. We used only one of them, in addition to the core SNePS: SNaLPS, the SNePS Natural Language Processing System. SNaLPS consists of a morphological analyzer, a morphological synthesizer, and a Generalized Augmented Transition Network (GATN) grammar interpreter/compiler (Shapiro 1982). The GATN grammar allows us to parse and generate natural language.

3.3 Modules of the algorithm

As mentioned in section 1.1 of chapter 1, the present work is based on our belief that English and French differ in their use of the three basic topological prepositions and not simply in a few idiosyncratic cases. On the contrary, we believe that there is more than simply convention involved in the use of topological prepositions in English and

in French. We believe, like Grimaud (1988), that metaphorical conceptualizations can explain the uses of these prepositions. The algorithm we propose implements this idea.

In this section, we describe the logical structure of the algorithm. We begin with a general overview, and then present it in more detail.

3.3.1 The general structure

The algorithm is divided into four phases:

Phase 1: Initialization

Phase 2: Derivation of the objective meaning of the sentence

Phase 3: Derivation of the target-language preposition

Phase 4: Finalization

The information used by the algorithm is of three different types:

Lexical information: Each word is stored in the lexicon together with relevant syntactic and morphological bilingual information about it.

Procedural information: Semantic information about each preposition, and bilingual conceptual information about objects, are both stored in procedural form.

Declarative knowledge: World knowledge about each object is stored in the knowledge base, made of SNePS networks.

Lexical information and world knowledge are two categories of knowledge that are used in several places in the program. They need to be stored in a way in which they are accessible at any time during the run of the program. This is the reason why they were stored in declarative form (the knowledge in the lexicon is also declarative). Semantic information about the prepositions and bilingual conceptual information about objects are both needed only at one place in the program and are therefore stored in procedural form, for the sake of reducing the amount of storage space needed by the program. The knowledge base is extended during the run of the program, when other types of knowledge are built that have to do with the sentence processed. The entities used in the knowledge base and in the lexicon will be described in section 3.4, along with the other entities built by the system.

The remainder of this section will be devoted to discussing the internal organization of each of the four phases of the algorithm. Each of these phases will be described in terms of the steps constituting them. Each step will have a module associated with it, which is the executing agent of that step. In this section, the input and output of these modules will be discussed. In addition, the types of knowledge they use, if any, will be mentioned. However, the details concerning the ways in which the modules fulfill their

roles will not be discussed here, but in section 3.5[1], after the main entities of the system are described.

3.3.2 Initialization

We will describe the steps of this phase and the types of information derived in each step, and identify which of the later phases need this information. The initialization phase is divided into two steps:

Step 1: Parsing the source sentence

Step 2: Building the conceptualizations of the located and reference objects

The module which performs Step 1 is called the Source Language Understander (SLU). The SLU takes as input the source sentence or noun phrase, and achieves three different goals:

1. It checks for lexical errors in the source sentence (that is it checks whether or not all the words used belong to the lexicon),
2. It finds the French translations of the words in the source sentence, with the exception of the preposition, and also the French syntactic and morphological information needed to generate the target language sentence.
3. It identifies the located object, the reference object, and the preposition of the sentence.

The SLU uses the syntactic and morphological knowledge stored in the lexicon.

The module which performs Step 2 is called the Conceptualization of Objects Builder (COB). The COB takes as input the located and reference objects, which were identified in the previous step, and builds all possible conceptualizations of these two objects; these are the conceptualizations that are possible regardless of the sentences in which these objects might appear, of their role in that sentence, and of which language the sentence is in. The COB uses some of the syntactic and morphological knowledge as well as world knowledge about the located and the reference objects, available in the lexicon and knowledge base, respectively. The COB is central to the algorithm. It will not be discussed in section 3.5 as the other modules are, but will be described in great detail in chapter 4.

The objects built or identified in the initialization phase are needed in the next three phases. Phase 2 uses the English conceptualizations built by the COB and the preposition identified by the parser module, phase 3 uses the French conceptualizations built by the COB, and phase 4 uses the French translations, the syntactic information, and the morphological information that was derived by the parser module.

[1]Actually, one module is not described in section 3.5; instead, it is the topic of an entire chapter, chapter 4.

3.3.3 Derivation of the objective meaning of the sentence

This phase takes as input the located object, the reference object, the preposition identified previously, as well as the English conceptualizations of the located and the reference objects, and builds the representations of the objective meanings of the sentence. The objective meanings of the sentence are not important in itself for our algorithm. What are important are the fact that they could be built, and the conceptual information that is related to them. The derivation of the objective meaning phase is divided into two steps:

Step 1: The identification of the English conceptualizations relevant to the sentence

Step 2: The building of the objective meaning

The module which performs Step 1 is called the Discriminator of English Conceptualizations (DEC). The DEC takes as input the preposition that was identified in the previous phase, and all the English conceptualizations that were built in the previous phase. It selects the English conceptualizations that are relevant to the sentence, given the preposition used. The DEC uses the semantic information about the topological prepositions, available from some part of the procedural knowledge of the system.

The module which performs Step 2 is called the Scene Builder (SB). The SB takes as input the English conceptualizations selected in the previous step and the located and reference objects identified in phase 1, and builds possible objective meanings of the sentence. As well, this module attaches to the objective meanings of the sentences the English conceptualizations associated with these objective meanings. This last operation is not needed in the process of translation, but will allow the system, at a later time, to compare the English and French conceptualizations of the same objective meaning and to issue a message stating whether these conceptualizations are the same in both languages or different. The SB uses world knowledge about the located and the reference objects, available from the knowledge base.

The output of the entire phase is the output of the SB, which will be described in section 3.4. This output will be used by the next phase.

3.3.4 Derivation of the target-language preposition

This phase takes as input the objective meaning of the sentence derived from the last phase, and the French conceptualizations of the located and the reference objects. It outputs the French preposition that is to be used in the French sentence or noun phrase. This phase is achieved in two steps:

Step 1: Derivation of the French conceptualizations relevant to the sentence

Step 2: Derivation of the corresponding French preposition

The module which performs Step 1 is called the MATCHER. The MATCHER takes as input the French conceptualizations of the reference object, which were output by the COB,

and one of the objective meanings of the sentence, which was output by the SB, and finds the French conceptualization corresponding to that objective meaning by matching pieces of their representations. It is in the MATCHER that the French and the English conceptualizations associated with the same objective meaning of a sentence are compared, and a message is issued. The MATCHER does not use any information directly from the knowledge base, lexicon, or procedural knowledge.

The module that performs Step 2 is called the French Preposition Deriver (FPD). The FPD takes as input the French conceptualization output by the previous module, and uses some of the syntactic and morphological information derived in phase 1. It outputs a French preposition. The FPD does not use any information from the knowledge base, lexicon, or procedural knowledge.

3.3.5 Finalization

This phase takes as input the French preposition output by phase 3 and the French translations of the English words used in the sentence (together with morphological and syntactical information about them), that were derived in phase 1. It generates the French sentence. The module of this phase is called the Target Language Generator (TLG). It outputs a valid French sentence or noun phrase translating the source sentence that was input to the system. No information from the knowledge base, the lexicon, or the procedural knowledge is used by the TLG either.

Figure 3.1 illustrates the organization of the algorithm that was just discussed.

3.3.6 General remarks

We want to emphasize two important points about our algorithm, and then add an important remark.

The first important point is that the knowledge used by the system — morphological, syntactic, semantic, world-based, and conceptual — seems to be knowledge human beings have access to. Our only point of hesitation has to do with conceptual knowledge: human beings may use more general conceptual principles and fewer conceptual rules needed for the process. We have tried to reduce the number of rules in a few cases by generalizing them. For example, we considered objects of the types of 'shade', 'fog', and 'picture' to have the same conceptual rule apply to them in English (this is not true in French) on the basis that these three objects create environments (see sections 4.2.1.13, 4.2.6.6, and 4.2.6.7 of chapter 4). These generalizations, however, are not backed with psychological data. More research in psychology has to be made, before general conceptual principles that would allow a reduction of the number of conceptual rules can be found. The conceptual knowledge we currently use will be described in chapter 4. Future work concerning this problem is discussed in chapter 7.

The second important point to emphasize is the claim that was already made in section 1.6 of chapter 1. This claim is that the located object is always conceptualized

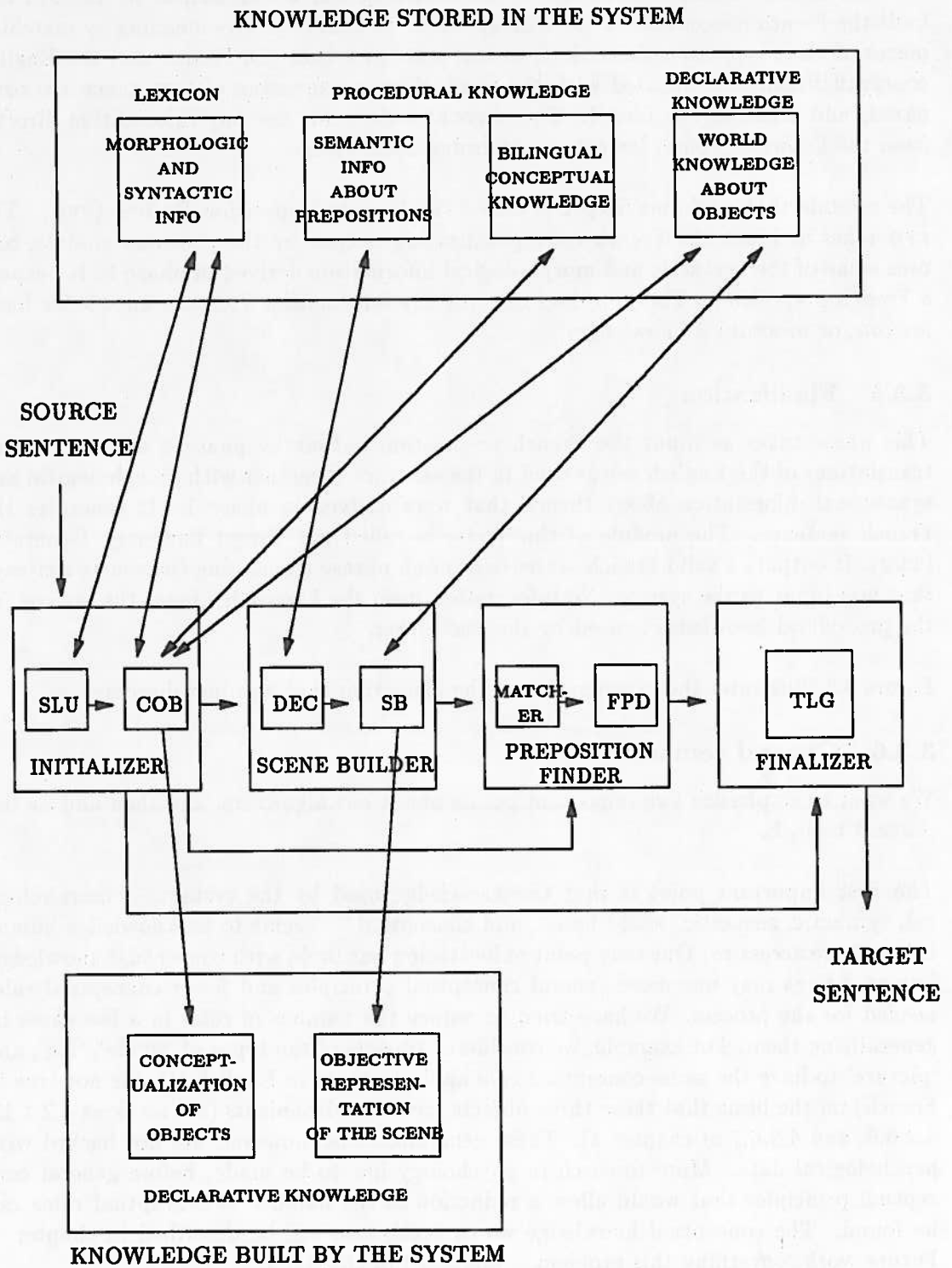


Figure 3.1: The structure of the algorithm

as a point, both in English and in French. Our conceptualizations of the scenes are therefore always of the form

"POINT ideal-meaning geometric-entity"

The ideal-meaning is one of the ideal meanings of 'in', 'on', or 'at', and the geometric-entity is one of: POINT, LINE, SURFACE, FULL VOLUME, EMPTY VOLUME, WATER-LINE, SOLID-LINE, OPPOSING-FORCE, and FUNCTIONAL CONTAINMENT.[2] Because the conceptualization of the located object does not influence the uses of the basic topological prepositions in English or in French (physical, sociological, and other properties about the located objects do, however, influence the uses), we decided to conceptualize every located object, whatever it is, as a POINT. The fact that we chose POINT as the conceptualization of the located object does not matter in itself. We thought that POINT would be the most neutral conceptualization, because any object can be seen as a POINT.

Finally, the remark we wanted to make is that, for the sake of clarity in our description of the algorithm, we did not discuss two extra features of our implementation: the treatment of errors in the source sentence and the treatment of ambiguity of the source sentence. These two features will be mentioned in section 3.5 and discussed in great detail in chapter 5.

3.4 The entities used in our theory

This section describes the main entities used in our algorithm. Such a description will be needed in the detailed explanation of the functioning of each module that we will give in section 3.5 and in chapter 4. We discuss two types of entities: the main entities already stored in the system before it runs, and the main entities actually built by the system. We do not discuss the conceptual rules stored in the system, because they will be discussed in great detail in chapter 4. Altogether, we discuss five groups of entities in this section:

1. The syntactic and morphological entities
2. The semantic representation of the prepositions
3. The world-knowledge representation of objects
4. The conceptual representations of objects[3]
5. The objective representations of a meaning of a sentence

[2] We define and discuss the ideal meanings of the prepositions and the geometric entities in the following section.

[3] Note that conceptual *representations* of objects are discussed in this section. It is the conceptual *rules* used to derive these conceptual representations that are not discussed here but are the topic of chapter 4.

The first three groups of entities are entities that are stored in the system, and the last two, are actually built by the system. Entities of type 1 are entities of the lexicon, entities of type 2 are entities of the procedural information, entities of type 3 are entities of the knowledge base, entities of type 4 are built by the COB and stored in the knowledge base, and entities of type 5 are built by the SB and stored in the knowledge base. We will now discuss each of these five entities.

3.4.1 The morphological and syntactic entities

This type of entity represents objects and relations in morphological and syntactic forms. These entities are, in fact, words, and information about these words.

The type of words used in the system are articles, nouns, verbs, and prepositions. The nouns represent material or immaterial objects and the prepositions specify spatial relations between two objects. The only verb allowed by our system is the verb 'to be', and it is not considered to have any meaning (and might as well be omitted in expressions, giving rise to the second type of expressions allowed by the system — noun phrases (see section 1.5 of chapter 1)). All the words used in our system are entered in the lexicon. The nouns, articles, and verbs are entered in English and French, but the prepositions are entered in English only; it is the algorithm's task to decide how the preposition should be translated into French. Each English word is entered in the lexicon, with the following information:

- its syntactic category: article, noun, verb, etc.,
- optional information about its syntactic or morphological irregularities,
- its tense and person, if it is a verb,
- its French translation (one English entry might have two French translations (see chapter 5)),
- its French gender.

The lexicon was built using some of the tools coming with SNePS that were described in section 3.2. To the usual information present in a lexicon, we simply added the specific type of information that is needed for translation between English and French. The reason why we limited our lexicon to the verb 'to be' or no verb at all is because we concentrate on the use of the basic topological prepositions. As Herskovits mentions, when a locative expression occurs in a case frame of a verb, frequently, depending on the verb, special conditions constrain the use of certain prepositions (Herskovits (1986) p. 127). Our study is not concerned with this type of case. As well, most of the verbs involve movement, which is not taken into consideration in our study, since we chose to concentrate on locative prepositions only. The reason why we did not add adjectives or the possibility of parsing more complicated sentence structures is that it would not have added any case of much relevance to our particular study.

3.4.2 Semantic representations of the prepositions

The semantic representations of a preposition is its ideal meanings: the set of prototypical uses of that preposition (see section 2.2.1 of chapter 2). Each preposition is represented by one or several ideal meaning(s). All the other uses of that preposition are uses that can be derived from the ideal meaning by applying conceptualization principles to the objects involved in the spatial relation.

The following is a list of the ideal meanings of each of the three basic topological prepositions, as defined in our system. These definitions are different from those of Annette Herskovits, who has one ideal meaning for each preposition.

AT: – Object contiguous with a point

ON: – Object contiguous with a line, or supported by a line
– Object contiguous with a surface, or supported by a surface

IN: – Object contained in a surface
– Object part of a full volume
– Object contained in an empty volume

When one preposition has multiple ideal meanings, these meanings are related, so we could have used a unique definition. However, we determined that these meanings are different enough to motivate the use of different prepositions in languages other than French and English, so we decided to distinguish them. This was done for theoretical consistency only. Note, however, that support and contiguity are not distinguished from one another. This is because we found that the distinction is too fuzzy: on the one hand, support of the type accounted for by the preposition 'on' is always accompanied by contiguity, and, on the other hand, it is very difficult to assess what type of contiguity does not involve support. We acknowledge, however, the fact that support and contiguity could be distinguished.

As was mentioned in section 3.2.1, the information about the ideal meaning is not declarative but procedural. This decision was made because this type of information is needed in only one module, and a little bit of storage space was saved this way.

3.4.3 World knowledge representation of objects

Another type of entity represents the world knowledge about objects, useful for our specific application. This information is objective in that it is language-independent, and could be used in applications other than ours, although it would probably have to be complemented for other applications. The entities representing world knowledge are represented by a SNePS network.

The world-knowledge base is organized following Sandra Peters's organization (Peters et al. 1987ab; 1988), based on work in cognitive psychology on categorization principles (Rosch et al. (1976; 1978), Mervis & Rosch (1981), Smith & Medin (1981), etc.). There

are three hierarchical levels: a superordinate level, a subordinate level, and a basic level. Objects can be defined at either the basic level or at a subordinate level. Properties are defined at the superordinate levels and at the basic level. The objects defined at the basic level, in addition to the properties stored at their level, inherit the properties of the superordinate levels. The objects defined at the subordinate levels inherit the properties of the superordinate levels and these of the basic level. We used Sandra Peters' organization for the representation of the world knowledge needed in our system. However, we adapted her representation according to our needs. In particular, we did not include all the superordinate levels that should be present in a hierarchy of her type because too great a number of these additional levels slowed our system down, without providing any information necessary for the processing of sentences related to our particular study. As well, we did not use her exact same representation: Peter uses default rules to represent basic-level properties, we do not so (see figure 3.3).

For the contents of the world knowledge base, we were inspired by the work of Kathleen Dahlgren (1988). Dahlgren's theory is called "Naive Semantics". It supports the idea of representing word meanings as common-sense knowledge, which is a set of naive beliefs about the structuring of the world. These naive beliefs might at times be vague, and even inaccurate. We did not concentrate all our effort on the design of a representation of world knowledge. We were looking for one that suited our purposes. Our main purpose was to be able to specify certain properties of the objects we consider, such as their shape, parts, function (in the sense of how human beings interact with them), and some other properties[4]. Dahlgren's organization gave us the freedom to incorporate all the properties that we needed to specify. Following Dahlgren, and inspired by her, some of the properties stored in the superordinate levels are shown in figure 3.2. In addition, we list the properties not shown in figure 3.2 because of lack of space: HUMAN, PART-OF-NATURE, BUILDING, OPEN-SPACE, CONTINENT, CITY, REFERENCE, PLACE, WATER-BODY, PART-OF-BODY, REPRESENTATION, FUNCTIONAL, PLACE-OF-WORK, INSTITUTION, PART-OF-BUILDING, KITCHEN-INSTRUMENT, PIECE-OF-FURNITURE, CLOTHING. The properties stored at the basic-levels are: SIZE, WEIGHT, SHAPE, MATERIAL, ADHERENCE, SUPPORT, HERMETICITY, MATERIALITY, NATURAL-FUNCTION, SOCIAL-FUNCTION, CAN-PERFORM. We will now explain the properties mentioned. We will explain only the non-obvious ones. We will first discuss the superordinate-level properties, and then will talk about the basic-level ones.

We start with the superordinate-level properties. Some of the distinctions made are useful for the study of the topological prepositions. Others are not but were made simply because they seemed relevant to the way people categorize objects. In the list we will just indicate which distinctions were needed for our specific study, and which were simply thought to be relevant to the way people categorize objects. The explanation of the importance of some of the properties will be discussed in the next chapter where all the uses that we considered will be explained.

[4]All the properties needed were described in section 1.3 of chapter 1.

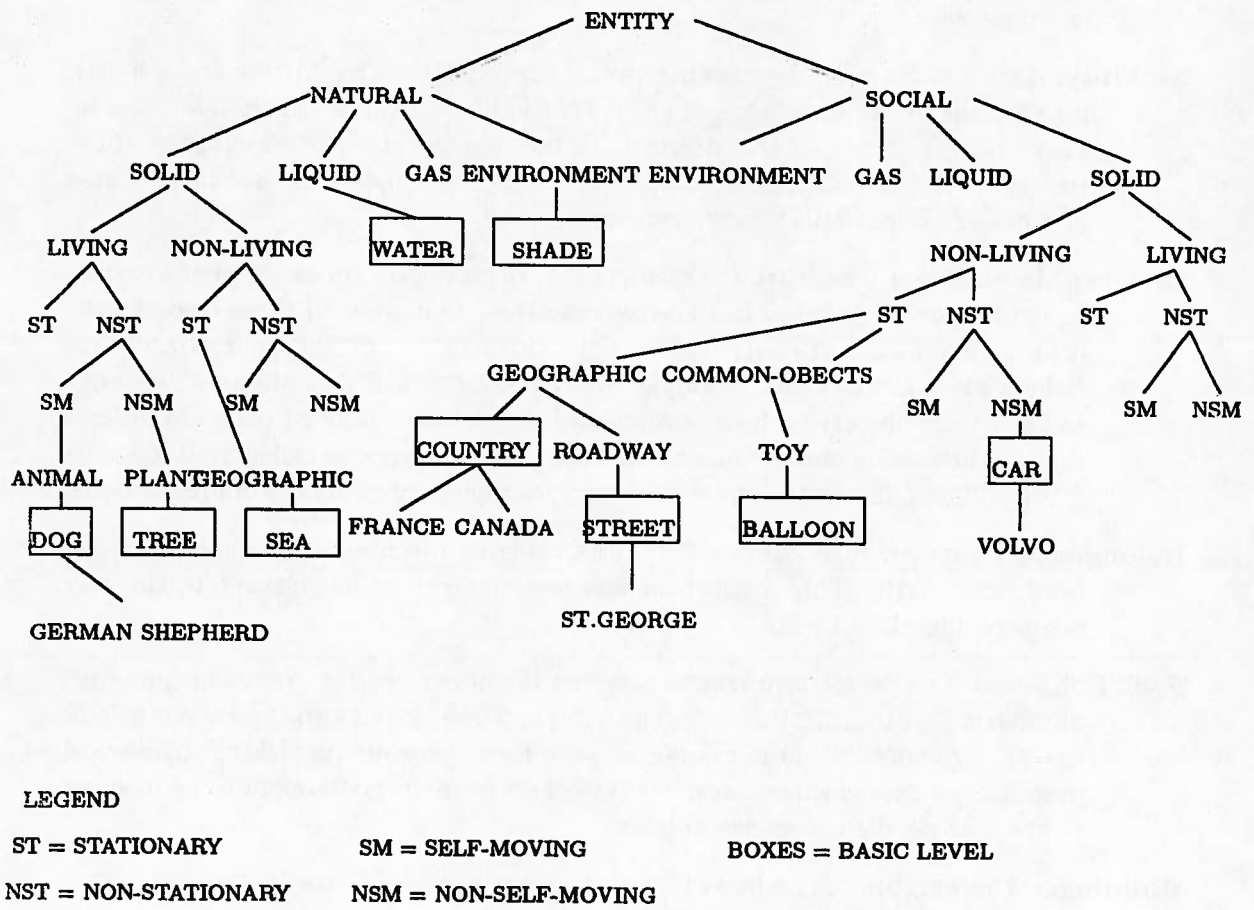


Figure 3.2: The organization of the world knowledge base

Entities: Every object, material or not, is an ENTITY. An ENTITY can be NATURAL or SOCIAL. An ENTITY is NATURAL if it is not one of man's creations, for example, animals, lakes, clouds, and rocks. An ENTITY is SOCIAL if it was created by man, for example, names of countries, buildings, cars, and knives. This distinction was simply thought to be relevant to the way people categorize objects.

Physical properties: An ENTITY can be SOLID, LIQUID, GASEOUS, or ATMOSPHERIC. ATMOSPHERIC ENTITIES are ENTITIES such as shade, the sun, the wind, and the middle (of the room). These distinctions were important to our particular study, in some cases.

Mobility: Every ENTITY can be STATIONARY or NON-STATIONARY. A tree and a house, for example are STATIONARY. If an ENTITY is NON-STATIONARY, it can either be SELF-MOVING or NON-SELF-MOVING. A boy and a cat, for example, are SELF-MOVING, while a car and a boat, for example, are NON-SELF-MOVING. These distinctions were important in some cases.

Geographic entities: These are the COUNTRIES, CITIES, CONTINENTS, REFERENCES, PLACES, WATER-BODIES, and ROADWAYS. Note that some of these GEOGRAPHICAL ENTITIES are SOCIAL, for example, COUNTRIES, CITIES, and ROADWAYS. Others are NATURAL, for example, WATER-BODIES and mountains. The REFERENCES are objects such as borders and the equator. The PLACES are objects such as the seaside and the mountains, i.e., *generic places*, as called by Herskovits (1986). These distinctions were necessary for the study of the use of prepositions.

Belonging to nature: The PARTS OF NATURE category includes objects such as rocks, sand, and earth. This distinction was just thought to be relevant to the way people categorize objects.

Social objects: The COMMON OBJECTS category includes REPRESENTATIONS (pictures, photographs, etc.), KITCHEN INSTRUMENTS, TOYS, PIECES OF FURNITURE. and CLOTHING. Some of these categories were needed for distinguishing topological prepositions uses in some cases, but the others were simply thought to be relevant to the way people categorize objects.

Buildings: The BUILDINGS can be FUNCTIONAL (stores, schools), DWELLINGS, PLACES-OF-WORK, and INSTITUTIONS. These distinctions were needed in the study of the uses of prepositions.

We will now describe the basic level properties. All the basic level properties that were incorporated are used in the study of the use of topological prepositions. Their importance will also be explained in the next chapter. Figure 3.3 illustrates our representation of this knowledge about trees. Node $M11$ represents the concept of the class of trees. If a LEX arc points from node i to node j, then j is a word that expresses the concept represented by i (Shapiro & Rapaport 1987). Thus, in Figure 3.3, the concept represented by node $M11$ can be expressed as 'tree'. Only English lexical information is included in these representations; as mentioned earlier, the French translation of an English word is stored with the English word in the lexicon.

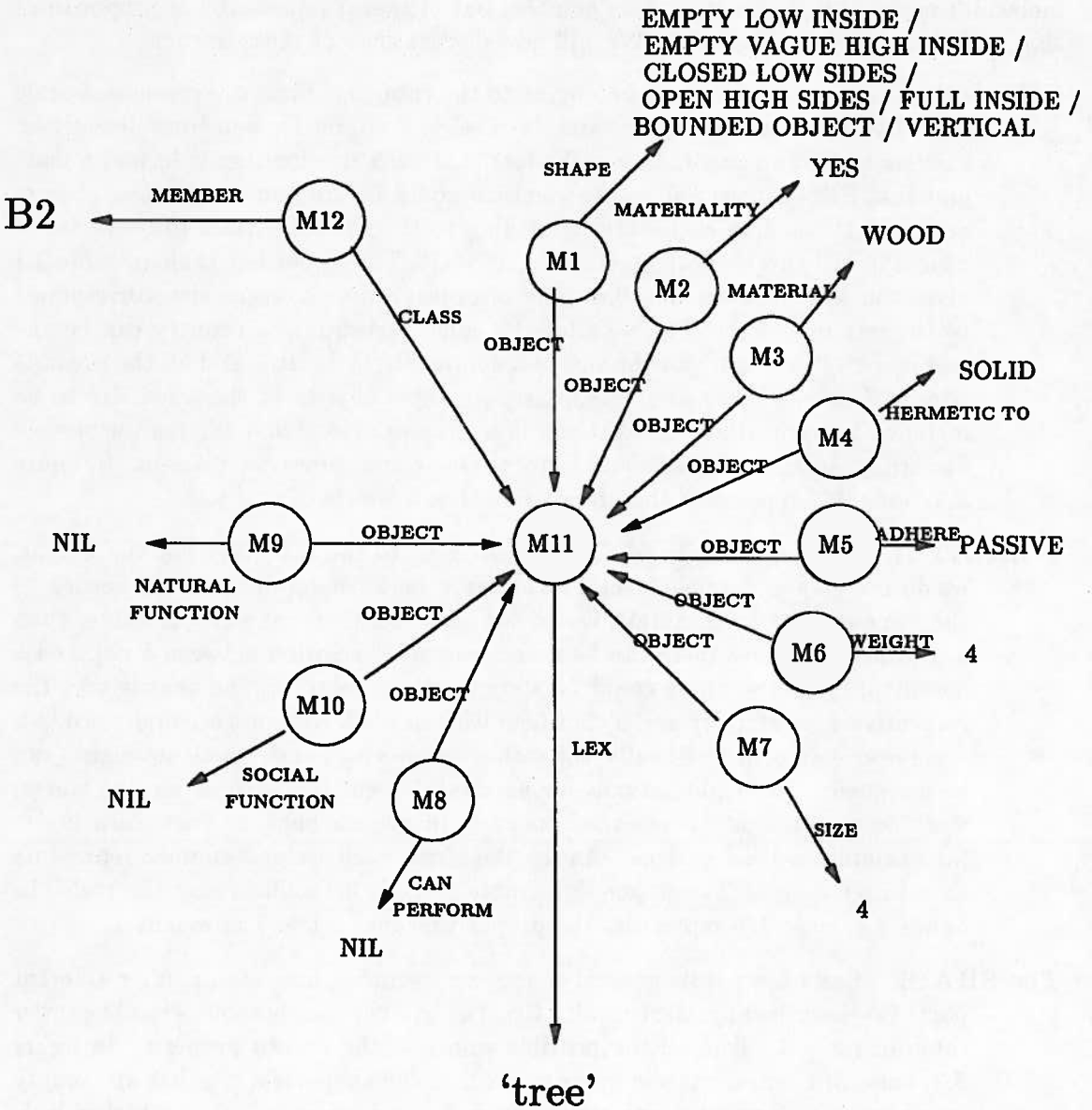


Figure 3.3: The basic level of the word 'tree'

Node $M12$ represents the proposition that the entity represented by node $B2$ is a member of the class of trees (Peters & Shapiro 1987ab, Shapiro & Rapaport 1987). The remaining molecular nodes (i.e., nodes with arcs pointing out of them) represent the propositions that trees have certain properties. We will now discuss each of these in turn.

SIZE is the average size of objects belonging to the category. Sizes are given on a scale from 1 to 11. In addition, the value "nosize" is available for uncountable entities, such as water and sugar. Given the fact that each size number is in fact a node and that SNePS does not create identical nodes in the knowledge base, objects are in fact linked to each other according to their relative sizes (objects are of course linked through other relations, as well). The upper left table in table 3.1 gives the size numbers and one type of object whose average size corresponds to the number. Note that we allow for some variations: a country can be the size of a city (Vatican), or the size of a continent (U.S.S.R.), and all the possible sizes in between these extremes. Also, we allow objects of the same size to be included in each other (the Vatican is in Rome). Note that for the purpose of our study, such a scale was satisfactory. One could, however, refine it. In figure 3.3, node $M7$ represents the proposition that a tree is of size 4.

WEIGHT is the average weight of objects belonging to the category. For the weight, we do not give as detailed a scale as for sizes, because details were not needed to the same extent. For example, we do not need to know that a city is lighter than a continent, because there can be no real support[5] relation between a city and a continent. However there could be a containment relation, and that is why the respective sizes of a city and a continent were needed. Weights are organized in a scale from 1 to 5. Additionally, the values "noweight", and "seemsnoweight" can be assigned. "Noweight" stands for uncountable entities such as air and water, and "seemsnoweight" is assigned to parts of objects such as parts of a body: for example, a head or legs. As for the sizes, each weight number represents an average weight. The upper right table in table 3.1 summarizes the scale. In figure 3.3, node $M6$ represents the proposition that a tree has weight 4.

The SHAPE of an object is its general shape, e.g., volume, line, etc., and its different parts (as described by Mark et al. (1987)), e.g., its sides, bottom, etc. The lower table in table 3.1 lists all the possible values of the SHAPE property. In figure 3.3, node $M1$ represents the proposition that the shape of a tree has an "empty low inside", an "empty vague high inside", "closed low sides", "open high sides", a "full inside", and that it is a "bounded object", and is "vertical". In reality, our representation is slightly different: there are actually as many shape arcs as properties about shape. In figure 3.3, only one shape arc is shown for the sake of brevity.

MATERIAL is the material an object is made of. The possible values of the MATERIAL property are: wood, rock, metal, glass, paper, and cloth. In figure 3.3, node $M3$ represents the proposition that a tree is made of wood.

[5]That is, a support relation requiring the located object to be lighter than the reference object.

size	*corresponding object*
nosize	water
1	flower
2	dog
3	human being
4	tree
5	building
6	lake
7	seaside
8	sea
9	country
10	continent
11	planet

weight	*corresponding object*
noweight	water
seemsnoweight	head
1	flower
2	dog
3	human being
4	tree, bus, house
5	geographical entity and bigger

type of shape	*values*
Parts	bounded high outside surface bounded low outside surface bounded outside surface closed low sides closed high sides open low sides open high sides bounded low inside surface empty low inside empty vague low inside empty high inside empty vague high inside full inside tiny width
General shape	line non-horizontal bounded object

Table 3.1: Sizes, weights, and shapes

ADHERENCE is the type of adherence relation that an object can have with another. The type of adherence might be *passive*, if the object adheres to another only because of a foreign agent (like in "a lamp on the ceiling", where it is the nails and screws that cause adherence between the lamp and the ceiling), *intermediate*, if the adherence is caused partly by one of the two objects involved in the relation and partly by a foreign agent (like in "a picture on the wall" where the force of the picture against the wall contributes to the adherence relation, together with the nail holding the picture), and *active*, if the adherence is caused by one of the two or both objects involved in the relation (like in "the fly is on the ceiling", where it is the fly which causes the adherence).[6]. In figure 3.3, node $M5$ represents the proposition that a tree cannot cause adherence in an adherence relation.

HERMETICITY is the property of an object to contain a certain type of object (a glass, for example, can contain a liquid but not a gas[7]. A balloon can contain both). The possible values assumed by this property are: solid, liquid, gas. In figure 3.3, node $M4$ represents the proposition that a tree is hermetic to solids only.

MATERIALITY is the property of an object to be material or not (e.g., shade is non-material, but dog is material). The values taken by this property are: yes or no (or NIL), meaning that the object is material or non-material, respectively. In figure 3.3, node $M2$ represents the proposition that a tree is material.[8]

NATURAL-FUNCTION is the inherent natural functions associated with the object. Note that we listed only natural functions relevant to our study. A tree, for example, has, among others, the natural function of adding oxygen to the environment. However, knowing this function of a tree is not necessary in determining the right preposition to use in a locative sentence; therefore, we did not include such functions. The values NATURAL-FUNCTION can take are: "social function", "create environment", "can support", "can hide", and "through-way". The value "social function" means the object has social functions that are pointed to by the SOCIAL-FUNCTION arc, "create environment" is what the shade, and the sun do: they create an environment in which other objects can be located. In figure 3.3, node $M10$ represents the proposition that a tree does not have a natural function of relevance to our study.

SOCIAL-FUNCTION are social functions served by an object. Such functions are: "cross countries", "take vacations", "live", "shop", "school-study", "child-work", and "swim". In figure 3.3, node $M9$ represents the proposition that a tree does not have a social function of relevance to our study.

---

[6]This distinction was based on Vandeloise (1986) pp. 200-202.

[7]A glass can, in fact, contain a gas. Our system, however, simplified relations between elements and containers.

[8]Note that the properties "Material" and "Materiality" could be merged into a single property.

CAN-PERFORM is the type of activity the object can perform. The values that this property can take are exactly the same ones as the values of SOCIAL FUNCTION above.[9] In figure 3.3, node $M8$ represents the proposition that a tree cannot perform any activity of relevance to our study.

In this general organization, we tried to put the maximum information in the higher levels, for space reasons. We gave the basic level all the information that could not be organized at a higher level.

3.4.4 Conceptual representation of objects

In this section we do not discuss the rules used to build the conceptual representations of objects. They will be discussed in chapter 4. Here, we present the conceptual representations themselves. Each object has one or several conceptualizations, represented by SNePS node(s). Each conceptualization contains the following information: LANGUAGE, ROLE, LOCOBJCONDITION, HIGHLIGHTS, EXTRA CONSTRAINT, DIMENSION, FULLNESS, and WIDTH.

Figure 3.4 shows a conceptual representation of a tree. Node $M14$ represents a set of conditions, node $M15$ represents a conceptualization, and node $M13$ represents the proposition that the node at the head of the META arc represents a conceptualization of the object at the head of the OBJECT arc, given that the conditions represented by the node at head of the CONDITIONS arc hold. We will now describe the conditions and conceptualizations, beginning with the conditions.

Conditions

Following are the kinds of conditions that are included in conceptual representations of objects. In our representations, each is the name of an arc that emanates from a node such as $M14$ in Figure 3.4.

LANGUAGE The language condition specifies the language or languages in which the conceptualization exists (English, French, or both). In figure 3.4, the LANGUAGE arc emanating out of node $M14$ points to "English" and "French", which means that for the language condition to be satisfied, the language in which the sentence is considered must be French or English. As in the case of SHAPE earlier, in reality, our representation is slightly different: there are actually as many language arcs as satisfying languages. In figure 3.4, only one language arc is shown for the sake of brevity.

[9]The reason why the values of SOCIAL-FUNCTION and CAN-PERFORM are the same is that in the checking of certain conditions, the located objects must have some actions they CAN-PERFORM that must match the SOCIAL-FUNCTIONS that can be performed at certain reference objects.

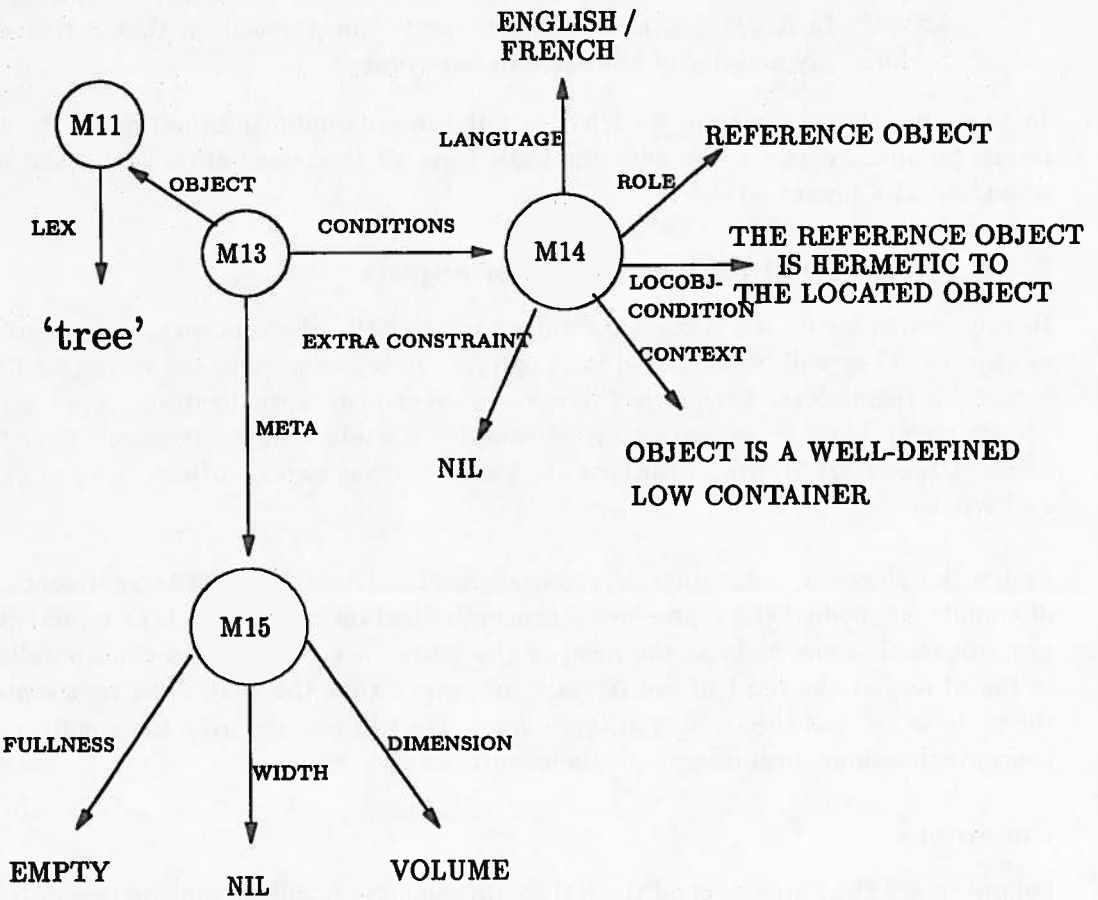


Figure 3.4: A conceptual representation of a tree

ROLE The role condition indicates the role that the object must have in the sentence — either a located or a reference object — for the conceptualization to be allowed. In figure 3.4, the ROLE arc emanating out from node $M14$ points to "Reference object", which means that for the role condition to be satisfied, the role of the object considered in the sentence considered must be "reference object".

LOCOBJCONDITION The locobjcondition condition names the conditions that the located object must satisfy alone or with respect to the reference object for the conceptualization to be valid (only objects whose role is to be a reference object might require such a condition). These conditions are related to world knowledge. In figure 3.4, the LOCOBJCONDITION arc emanating out from node $M14$, points to "The reference object is hermetic to the located object", which means that for the condition to be satisfied, the condition concerning the located object is that the reference object must be hermetic to the located object.

HIGHLIGHTS The highlights condition describes the parts of the object conceptualized and its properties that must be highlighted for this specific conceptualization to be instantiated. This field is central to the development of the program because it is the one that will be used to build the objective meaning of the sentence and to find a corresponding French conceptualization. In figure 3.4, the HIGHLIGHTS arc emanating out from node $M14$, points to "Object is a well-defined low container", which means that for the highlights condition to be satisfied, the fact about the reference object that must be highlighted is that its lower part is well defined.[10]

EXTRA CONSTRAINT The extra constraint condition (short for "extra-sentential constraint") is a constraint on the extra-sentential context in which the sentence is uttered for the conceptualization to be valid. Whatever the extra constraint condition is, the algorithm considers it to be verified, given that the algorithm processes isolated sentences. For this reason, it cannot know what the extra-sentential context is. The point of having the EXTRA CONSTRAINT condition is to have an indication of the additional conditions that our system, if transformed into a system which processes paragraphs, should verify together with the conditions already mentioned. In figure 3.4, the EXTRA CONSTRAINT arc emanating out from node $M14$, points to NIL, which means that there would be no EXTRA CONSTRAINT to be satisfied, even in the case where the system was extended to a system that processes paragraphs.

Conceptualizations

We will now describe the constituents of a conceptualization. Each is the name of an arc that emanates from a node such as $M15$ in figure 3.4.

[10]Note that HIGHLIGHT is not really a condition, since it does not need to be checked. It is rather a statement indicating that when a given conceptualization is taken into consideration, the highlighted property of the object being conceptualized is specified by the node pointed to by the HIGHLIGHT arc.

DIMENSION The dimension constituent is the type of conceptualization. Its value can be POINT, LINE, SURFACE, VOLUME, WATER-LINE, SOLID-LINE, OPPOSING-FORCE, or FUNCTIONAL-CONTAINMENT. It should be noted that the types of conceptualizations in which an object can be conceptualized are not simply geometric. All the English metaphors are geometric, but not all the French ones are. The non-geometric conceptualizations are: WATER-LINE, SOLID-LINE, OPPOSING-FORCE, AND FUNCTIONAL-CONTAINMENT. Each of these are related to geometrical concepts but are not exclusively geometrical.

FULLNESS can take the values FULL or EMPTY . This parameter determines whether we focus on what an empty volume contains or on what components the object is made of or on what can be embedded in it.

WIDTH is the width of the conceptualization. Its value can be TINY-WIDTH or NIL. NIL stands for "the conceptualized width does not matter". The only reason why we need this field is that objects of the category "representation" (picture, photographs, etc.) are conceptualized, according to us, as a volume with a tiny width in English, and as a surface in French. The conceptualized width does not matter in any other cases.

Discussion

Conceptual representations of objects are central to our algorithm. Our representation, illustrated in figure 3.4, reflects our understanding of Lakoff & Johnson's (1980) theory of metaphor, that is, given that certain conditions hold, a metaphor highlights certain aspects of an object and hides others. In fact, a metaphor views an object (O1) as if it was another simpler one (or a geometrical entity) (O2) whose main features (F) correspond to the highlighted features (F) of the first object (O1). We will now explain our motivations for the building of each feature of the entity representing a conceptualization of an object. The motivation for the entire entity comes from Lakoff & Johnson (1980) and Grimaud (1988). The type of information pointed to by the META arc is in some respect suggested by parts of Clark (1973), in particular, the DIMENSION field: as mentioned in section 2.4.1, Clark (1973) characterizes 'at', 'on', and 'in', as uni-, bi-, and tri- dimensional. We use this characterization in our geometrical conceptualizations where we consider a point to be uni-dimensional, a line or a surface to be two-dimensional, and a volume to be tri-dimensional. However, this sole geometric conceptualization was not enough: we needed to add some information to the conceptual representations in order to use them to translate local prepositions from English to French. First, as mentioned above, extra types of conceptualizations, not necessarily purely geometrical, had to be added, because French uses some non-geometrical conceptualizations in some cases. Second, we needed to add the FULLNESS field, since we decided to differentiate the conceptualizations where objects are seen as an entity, or as a set of components. Third, we added the WIDTH field because, as mentioned above, an object such as a picture, or a photograph is conceptualized as having a width in English, but not in French, and this makes a difference in the use of prepositions. For example:

(56) The man *in* the picture.

is translated into French as

(57) L'homme *sur* la photo.
(The man *on* the picture)

The type of information pointed to by the CONDITIONS arc was partly inspired by Grimaud, and partly derived by us. First, the LANGUAGE field was inspired by Grimaud who says, as we saw in chapters 1 and 2, that the conceptualizations of objects are different in English and French. That is, it may be possible to conceptualize an object in a certain way in one of the languages but not in the other. So, we need to require that the conceptualization pointed to by the META arc is a possible one for the given object in the language considered. Second, all the other fields pointed to by the CONDITIONS arc were deemed necessary as a result of our examination of the uses of the three basic topological prepositions.

The ROLE field was included because we believe that only the conceptualization of the reference object matters to the use of the three basic topological prepositions. We therefore had to make sure, in the sentence, that the conceptualization we are considering is a conceptualization of the reference object and not of the located object.

The LOCOBJCONDITION field imposes conditions on the located object or on the relation between the located and reference objects that are required for the metaphor pointed to by META to be instantiated. The LOCOBJCONDITION conditions are very important because even though the conceptualization of the located object does not have a determinant role in the use of the prepositions studied, some physical properties about the located object alone or in conjunction with the reference object are determinant factors in the use of these prepositions and must be considered. An example of a case where a condition about the located object is necessary in the use of 'in' and 'dans' is:

(58) * The oxygen is in the glass.
L'oxygène est dans le verre.[11]

where the sentences are not valid because oxygen is a gas, and a gas cannot be contained by a glass. All these conditions will be discussed in the next chapter.

The HIGHLIGHTS field is central to the system. It names the objective properties of the object considered, which are highlighted when this object can be conceptualized as it is described in the node pointed to by META. Because the HIGHLIGHTS field describes an objective situation, it might be common to English and French conceptualizations. In fact, this is the main point of the algorithm: the matching of the French conceptualization corresponding to the scene is done by finding a French conceptualization with the same HIGHLIGHTS field as the English conceptualization.

[11] Note again that sentences (58) are rejected by our system because of its simplification of the containment relation. In reality, such sentences are valid.

The EXTRA CONSTRAINT field is not necessary for our program. We included it simply because it was a nice extra feature, given that the EXTRA CONSTRAINT field accounts for problems that a system more sophisticated than ours would have to take into consideration.

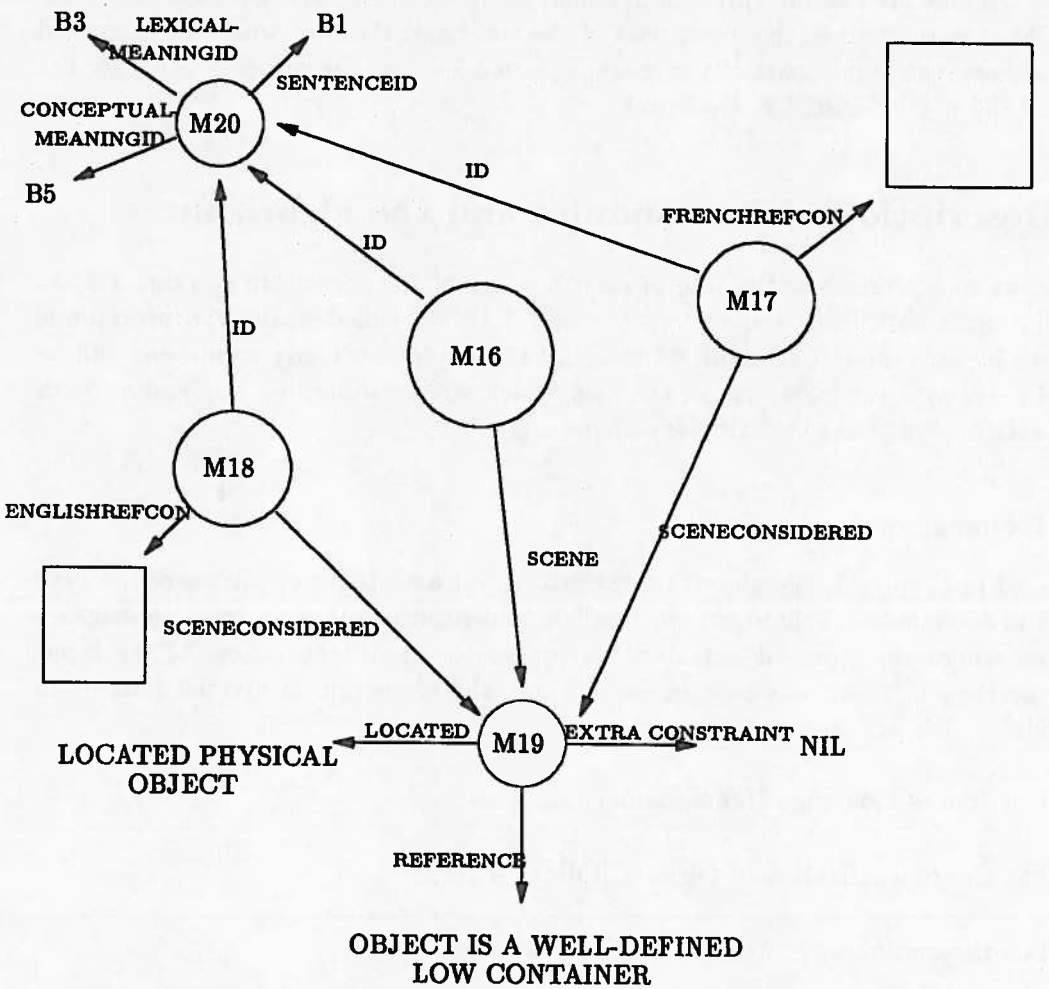
3.4.5 Objective representation of a meaning of a sentence

The entities that represent possible objective scenes that can be described by a particular sentence are language-independent. Figure 3.5 illustrates, among other things[12], our objective representation of the sentence "The boy is in the tree". Node $M20$ represents the identity of the sentence, node $M19$ represents an objective scene. Node $M16$ represents the proposition that the sentence of identity $M20$ is represented by an objective scene of the type represented by node $M19$.

Node $M20$ has three arcs emanating from it. The SENTENCEID arc indicates the identity of the English sentence being currently processed, the LEXICALMEANINGSID arc points to the identity of the combination of the meanings of the located and reference objects being processed, and the CONCEPTUALMEANINGID arc points to the identity of the scene being processed. The SENTENCEID arc distinguishes the current nodes of the run from the nodes that were built in a preceding run. The LEXICALMEANINGSID arc is needed in the case of *lexical ambiguity*: if an English noun has two French translations which call for a different use of prepositions, then some nodes are built in each of the two cases, and need to be differentiated. The LEXICALMEANINGSID arc makes it possible to differentiate them. Chapter 5 will explain this point more in detail. Finally, the CONCEPTUALMEANINGID arc is needed in the case of *conceptual ambiguity*: it allows to differentiate between possible objective scenes of one particular sentence. The same sentence might represent several objective scenes, and the nodes built in each case must be differentiated. Chapter 5 will discuss this as well. Node $M19$ has 3 arcs emanating from it. The LOCATED arc points to the objective information about the located object, the REFERENCE arc points to the objective information about the reference object and the EXTRA CONSTRAINT arc points to the extra-sentential constraint that should be satisfied (but cannot be checked by our system) in order for this scene to hold. The information pointed to by the LOCATED and REFERENCE arcs comes from the HIGHLIGHTS arc of the conditions for the conceptualization of the located and reference objects to be instantiated, respectively.[13] The EXTRA information is the information pointed to by the EXTRA CONSTRAINT arc in the conditions for the conceptualization of the reference object to hold. All this information characterizes the objective scene represented by the sentence. As mentioned in section 3.3.2, this entity in itself is not important to the system. What is of importance is its existence; therefore, we did not need to represent the scene in a very detailed way. Figure 3.5 does not show a scene alone, since it is not such an important entity; instead, it represents the scene together with its English and French

[12]The "other things" shown in Figure 3.5 are the conceptualizations added to the objective representation of a meaning of a scene.

[13]The highlighted information about the located object is always the same, since the same conceptual representation is used in all cases for the located object.



LEGEND:

= CONCEPTUALIZATION OF THE REFERENCE OBJECT INVOLVED IN THE SCE
(IN THIS PARTICULAR CASE, EACH OF THE TWO CONCEPTUALIZATIONS
IS EMPTY VOLUME)

Figure 3.5: The representation of an objective scene together with the English and French conceptualizations of the sentence:"The boy is in the tree"

conceptual information. We illustrate this representation with the sentence "The boy is in the tree". There are two interpretations of this sentence in English: the case where the boy is located in the trunk (the lower part of the tree) and the case where he is located in the branches (the higher part of the tree). Figure 3.5 shows the representation for the case where the boy is located in the trunk.

3.5 Description of the modules and the algorithm

Now that we have described the role of each module of the algorithm (section 3.3) as well as the main entities of the system (section 3.4), we will describe the procedures undertaken by each module to fulfill its role. All the modules already mentioned will be described here, with the exception of the COB which will be described in Chapter 4. In the final section, we present a summary of the algorithm.

3.5.1 General review

As discussed in chapter 1, the algorithm takes as input a sentence or phrase of the type described in section 1.5 of chapter 1, in English, and returns either an error message — in the case where an error was detected[14] — or one or more translations of the input sentence in French[15]. As was seen in section 3.3, the algorithm is divided into seven main modules. We list these modules again:

- The Source Language Understander (SLU)
- The Conceptualization of Objects Builder (COB)
- The Discriminator of English Conceptualizations (DEC)
- The Scene Builder (SB)
- The MATCHER
- The French Preposition Deriver (FPD)
- The Target Language Generator (TLG)

3.5.2 The modules

We will discuss the SLU, DEC, SB, MATCHER, FPD, and TLG in the order in which they are invoked by the algorithm.

[14]The various levels of error detection are discussed in chapter 5.

[15]The different types of ambiguity dealt with in the program are discussed in chapter 5.

The Source Language Understander

The functioning of the SLU is very basic. The SLU relies on the Morphological Analyzer and on the GATN Grammar of SNaLPS. The parts of the SLU built for this program are the lexicon and the parsing grammar. As mentioned in section 3.3.1, the SLU has three tasks: checking that each word in the sentence exists in the lexicon, finding the French translation of each word and information about these translations, and identifying the objects. The first task is done in a standard way; therefore, we will not describe it. The second task is simply a matter of fetching the right information in the lexicon, and the third task assigns roles according to the syntactic roles of the words in the sentence — the subject is the located object, the object is the reference object, and there is only one preposition. The SLU could easily be extended to handle more complicated sentences than the ones we considered.[16]

The Discriminator of English Conceptualizations

The DEC has a straightforward role: it must select all the English conceptualizations that could apply to the scene, according to the preposition used in the source sentence. To do this, it first decides, according to the preposition used, what the ideal meaning of the sentence could be. This operation is a simple matching operation: to each preposition corresponds a set of ideal meanings, as was discussed in section 3.4.2. The next step consists of selecting English conceptualizations. This can be done easily, since to each ideal meaning there corresponds one metaphor. Once the ideal meanings are selected, the DEC knows what metaphors it is looking for. It then simply needs to find, among all the English conceptualizations built by the COB, the conceptualizations whose DIMENSION arc points to information corresponding to that of one of the ideal meanings selected. Once this operation is done, the DEC simply marks the conceptualizations it found. Several alternative might arise: the DEC might select one, several, or no conceptualizations. These cases will be discussed in chapter 5.

The Scene Builder

The SB builds the set of the possible scenes expressed by the sentence, for a given pair of meanings of the located and reference object[17]. The SB works in two phases. In the first phase, it attempts to build the set of the possible scenes expressed by the sentence. In order for this set to be built, each candidate scene for this set has to fulfill two requirements which both depend on the relation between the located and the reference objects. The first requirement is imposed by the field LOCOBJCONDITION of the chosen conceptualization of the reference object. Each possible such requirement will be described in chapter 4. The second requirement is that the sentence must be normal.[18] In our application, we restrict the notion of normality to the normality of the relative sizes

[16]There was no point for us, however, to extend it since we were working on a definite problem which did not need such an extension.

[17]It is therefore the SB that detects complex lexical ambiguity (which will be discussed in chapter 5).

[18]See chapter 5 for a discussion of wether or not the two requirements discussed here should be distinct from one another.

of the objects (for example, "The elephant is on the flower" is not allowed in a normal context), and to the materiality of the located object, which is sometimes compulsory in a normal context (for example, "the sunshine is in the box" is not allowed in a normal context).[19] In its second phase, the SB actually builds the set of scenes that fulfilled the two requirements. These scenes are of the form described in section 3.4.5. As well, the SB links the English conceptualization of the reference object to the scene. In figure 3.5, node $M18$ is the link of the English conceptualization to the scene. $M18$ has three arcs emanating out from it: SCENECONSIDERED points to the scene, ENGLISHREFCON points to the English conceptualization of the reference object, and ID points to the identity node of the instance considered.

In the case of simple or complex conceptual ambiguity (which will be discussed in chapter 5), the MATCHER, FPD, and TLG are each called as many times as there are scenes built by the SB. In the following discussions, we describe the action of each module on a single scene, bearing in mind that this module might be called several times.

The Matcher

The MATCHER derives the French conceptualization that will be used for the French sentence, in the context considered. The MATCHER matches the scene considered to a French conceptualization. This is done by finding a French conceptualization of the reference object which has the field HIGHLIGHTS filled with the same information as the field REFERENCE of the objective scene. Once a conceptualization is found, it is marked and added to the objective scene of the sentence, thereby completing the picture that was represented in figure 3.4. In figure 3.5, node $M17$ is the link of the French conceptualization to the scene. $M17$ has three arcs emanating out from it: SCENECONSIDERED points to the scene, FRENCHREFCON points to the English conceptualization of the reference object, and ID points to the identity node of the instance considered.

The French Preposition Deriver

The FPD derives the French preposition to be used in the French sentence, given the scene considered. This is done simply by matching the French conceptualization found by the MATCHER to a list of French prepositions. The list of French conceptualizations together with the corresponding French preposition is kept in procedural form, again because this information is required only once. This list was not described in section 3.4.2 because it does not present any interesting feature: it is a simple straight-forward look-up table. For the sake of completeness of our description, this look-up table is shown here in table 3.2.

The Target Language Generator

Like the SLU, the TLG relies on the Morphological Analyzer and the GATN grammar of SNaLPS. The TLG takes as input all the nodes created by the SLU that represent French

[19] We use some of Herskovits's notions of normality, in this part.

Conceptualization	*French Preposition*
point	à/au
line	sur
surface	sur
bounded surface	dans
full volume	dans
empty volume	dans
water line	au bord de/du
solid line	en bordure de/du
opposing force	contre
functional containment	en

Table 3.2: The look-up table of the FPD

words or information about French, and the French preposition output by the FPD. The generator using this information generates a grammatically correct French sentence of the same form as the English one input by the user, for the scene considered. The only extra functions we needed to build for generation purpose were the morphological functions. The two morphological functions existing in SNePS were built for English generation only. We needed to build our own French morphological functions in order to generate French sentences.

3.5.3 The algorithm

We end this chapter with a summary of the algorithm. The two REPEAT-UNTIL loops in the algorithm are designed to deal with ambiguity. These loops will not be discussed until chapter 5. A general comment about them is that the inner loop is needed for simple and complex cases of conceptual ambiguity and the outer loop is needed for complex cases of lexical ambiguity (simple ones do not require a loop). The meaning of simple and complex cases of lexical or conceptual ambiguity will be given and discussed in chapter 5.

The second type of procedure in the algorithm that was not discussed in this chapter is the type that deals with error detection. All types of errors will also be discussed in chapter 5.

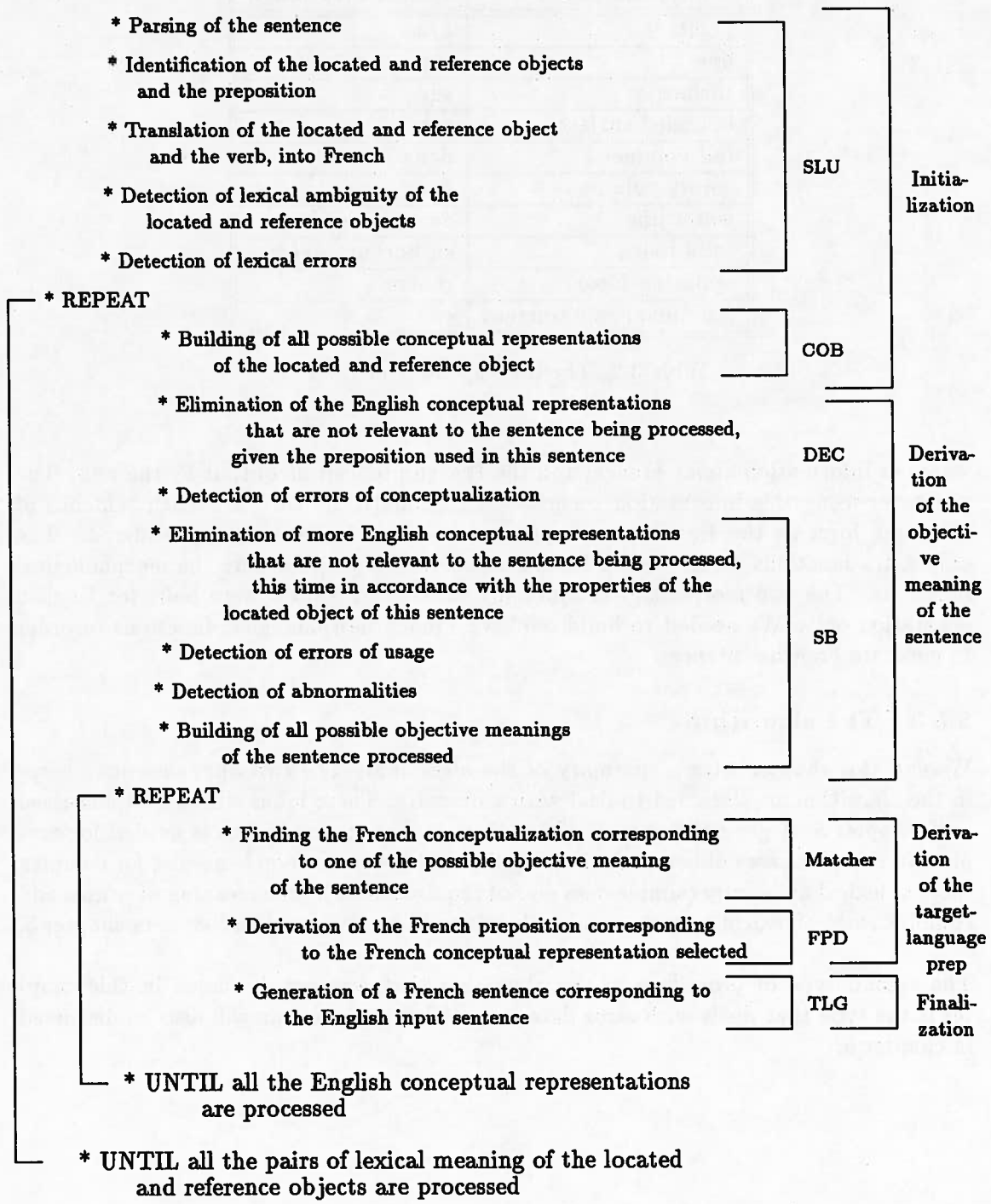


Figure 3.6: The outline of the algorithm

Chapter 4

Conceptualization of Objects Builder

This chapter discusses a very important module of the system: the Conceptualization of Objects Builder (COB). As mentioned in section 3.2.4, the COB builds conceptual information about objects during the execution of the program. Section 3.2.4 already discussed the output of the COB — the representation of the conceptual knowledge. We will now concentrate on the COB's actual building of this conceptual knowledge.

In the first section, we will overview the general structure and functioning of the COB. The next section will discuss the details of how the COB functions for each conceptual representation that can possibly be built by the system. In the last section, the differences between English and French will be summarized and analyzed.

4.1 Overview of the COB

The COB takes as input the name of one object and builds all possible conceptual representations of this object, regardless of language, situation in which it is involved, etc. The COB uses mainly world knowledge about the object, but also, in certain cases, some morphological and syntactic information about the name of that object. We will first concentrate on the general functioning of the COB, and then we will discuss the type of knowledge used by the COB during its execution.

4.1.1 The functioning of the COB

The COB is structured as a set of pattern-action rules. Each pattern-action rule is also called a case. In order to avoid confusion with other similar structures, we called each pattern a set of *Provisions* (or the provisional part of a case). The Provisions check whether certain world-knowledge properties about the object considered and certain morphological or syntactic properties about the word describing the object are true. In certain cases, however, they simply check the identity of the object considered. The exact type of information checked by the Provisions will be described in section 4.1.2. The exact set of Provisions used in each pattern-action rule of the COB will be discussed each

separately in section 4.2. The action rule corresponding to a set of Provisions (which we also call the execution part of a case) builds a conceptual representation of the type described in section 3.2.4. Figure 4.1 shows a conceptual representation of the object 'tree'. We recall what a conceptual representation is. A conceptual representation is an entity made of a conditional part and a metaphorical one. In order to avoid confusion with the overall pattern-action structure, the conditional part of each conceptual representation is called its set of *Conditions*. Conditions are different in nature from Provisions because they depend on the sentence in which the object being conceptualized is mentioned, whereas the Provisions depend on the properties of the object alone.[1] The set of Conditions checks whether its corresponding metaphor can be instantiated given the sentence considered. In particular, the Conditions check the following points about the object being conceptualized and the sentence considered:

- the role in the sentence of the object being conceptualized,
- the languages in which the sentence is uttered,
- the properties of the object being conceptualized that are highlighted,
- world-knowledge conditions involving the located object of the sentence,
- extra-sentential constraints.

If all the Conditions of a conceptual representation are satisfied, its metaphor can be instantiated. In figure 4.1 for example, if 'tree' is the reference object of the sentence, and the sentence is uttered in English or French, and the highlighted property of 'tree' is that it is a well defined low container[2], and the condition involving the located object is that the reference object is hermetic to the located object, then the empty volume metaphor can be instantiated.[3] The metaphorical part of a conceptual; representation is a description of the metaphor. It includes the following three types of information about the metaphor:

- its dimension,
- its fullness,
- its width.

In figure 4.1, the dimension of the metaphor is *volume*, its fullness is *empty*, and its width is irrelevant.

[1] The highlight Condition, however, is different from the others in that it is not a condition, but rather, a statement. It does not depend on the sentence, but instead, qualifies it. In addition, in certain cases it overlaps with a Provision.

[2] As mentioned previously, the highlight condition is particular: in fact, the highlight condition is not really a condition because it is always true. Rather than a condition, it is a piece of information specifying what properties of the object, its conceptualization highlights.

[3] Note that no extra-sentential condition needs to be satisfied in this case.

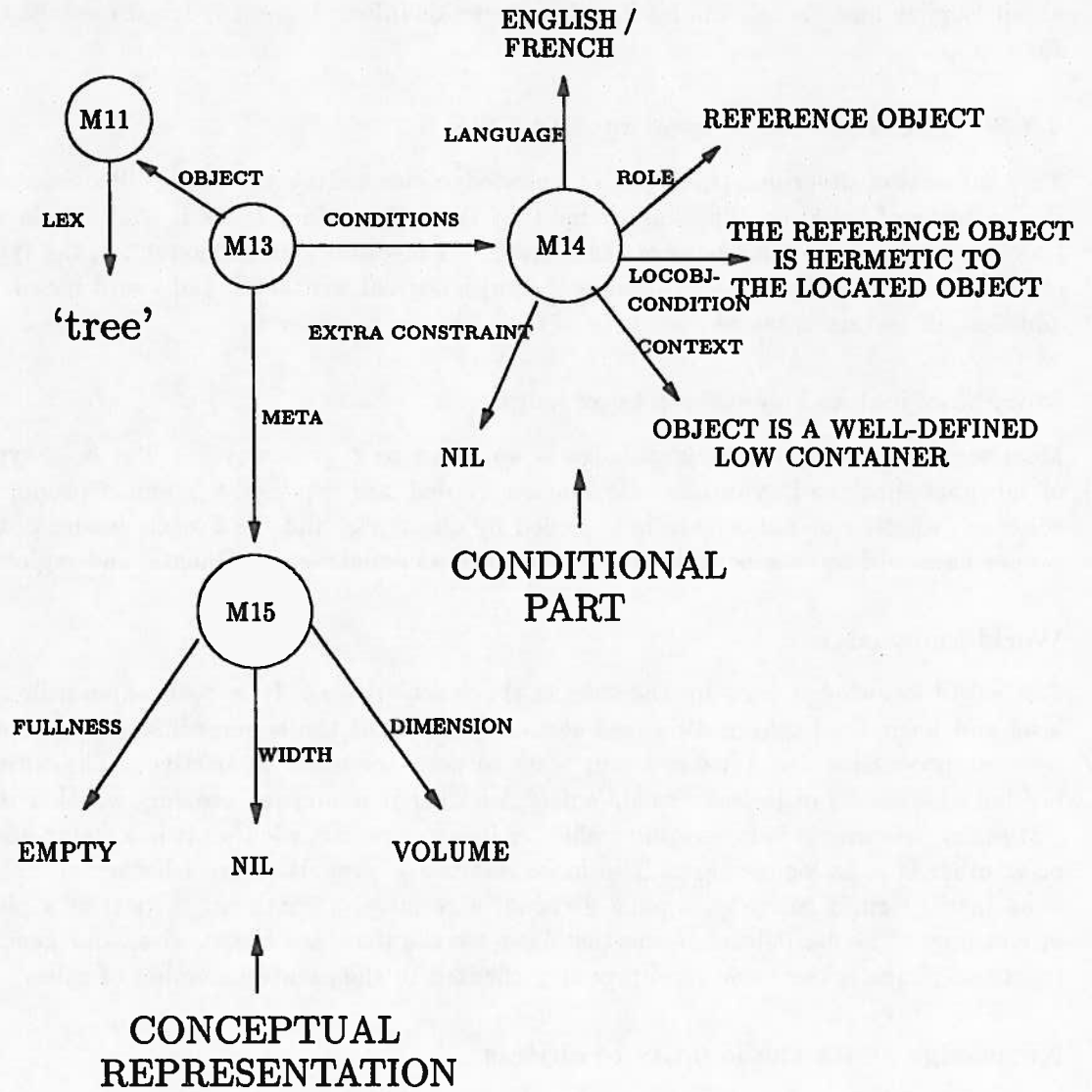


Figure 4.1: A conceptual representation of a tree

As was mentioned in section 3.2.4., we consider that a lot of external information has to be given to the COB. We would hope that in the future, more general conceptual rules about English and French will be found so that this information is reduced (see chapter 7).

4.1.2 The knowledge used by the COB

This subsection describes the types of knowledge checked by the sets of Provisions of the patterns of the COB. The knowledge that these Provisions check is available in the lexicon and the knowledge base of the system. As mentioned in section 4.1.1, the types of information used by these Provisions is morphological, syntactic, and world-based. In addition, in certain cases, the identity of the object is considered.

Morphological and syntactic knowledge

Morphological and syntactic knowledge is not used to a great extent. The only types of morphological and syntactic information needed are whether a noun is proper or common, whether or not a noun is preceded by an article, and the French gender of the proper names of some geographical entities, such as countries, continents, and regions.

World knowledge

The world knowledge used by the COB is the most critical. It is both superordinate-level and basic-level information (see section 3.2.3.). At the superordinate level, some relevant properties are quite general, while others are more restrictive. The former involve whether an object is a social entity, whether it is non-self-moving, whether it is stationary, whether it is geographic, whether it is geometric, whether it is a water-body, or whether it is an open-space. The more restrictive properties are whether an entity is an institution, a roadway, a page division, a country, a continent, a city, or a piece of clothing. The basic-level properties used by the COB are shape, size, and general function. Shape is the basic-level property checked in the greatest number of rules.

Knowledge about the identity of objects

In certain cases, the information needed about an object in order to know if a certain conceptualization is possible includes its identity. The COB considers three such cases in English: 'sea', as in Herskovits's case "Spatial entity at sea"; 'wall' as in Herskovits's case "Physical object contiguous with a wall", and 'air', as in Herskovits's case "Physical object 'in the air'" (Herskovits 1986, chapter 9).

4.2 Detailed functioning of the COB

This section concentrates on each of the conceptual representations that can be built by the COB. Each is described in a subsection. For each, the Provisions that must be satisfied for the conceptual representation to be built are shown first. This is followed by the set of Conditions of the conceptual representation built, that is, the conditions

that must be satisfied in order for the conceptualization (metaphor) of this conceptual representation to be instantiated. In addition, each of the sections highlights the cases in which the conceptual representation described is different in English and French, and finally, each conceptual representation is compared with the cases discussed in other research (Herskovits 1986; Vandeloise 1986; Mark 1990). The descriptions are grouped together according to the metaphorical part of their conceptual representation The information about the metaphorical part of each conceptual representation is described at the beginning of the section each of these cases belongs to.

The different groups of conceptual representations studied are these whose conceptualization is: a point, a line, a surface, a bounded surface, a full volume, an empty volume, an empty volume with a tiny width, a water body edge, a solid line, an opposing force, a functional container, an intersection, and an opposition. The name of each subsection describing the diverse conceptual representations is the highlighted property of the conceptual representation. The highlighted property of each conceptual representation is central to the system.

Most of the cases discussed here are adaptations of the cases discussed by Herskovits (1986 chapter 9). We, however, changed a few of these cases, because we found that they were not specific enough. Additionally, we discuss some new cases which deal with the French use of the prepositions studied.

4.2.1 Conceptualizations as a point

In each of the conceptual representations described in this section, the node pointed to by the arc META contains information specifying that the dimension[4] of the metaphor is POINT, and that its fullness and width are NIL. This conceptualization is a geometric one, and indicates that the object considered is conceptualized as a point.

The first conceptual representation described in this section is the unique conceptual representation of our system that applies to the located object (see chapter 3). All the consecutive ones only apply to reference objects.

4.2.1.1 Object is a (physical) located object

As was mentioned in section 3.3.6, any object is conceptualized as a point, when used as a located object. Because any object can be a located object, the provisional part of the case is empty.

In the execution part of the case, the conditions of the conceptual representation are the following:

[4]The term 'dimension' is misleading here, because it gives the reader the impression that the dimension arc points to numerical information, which it does not. The dimension arc points to geometrical information (point, surface, line, etc.) or to functional information (obstruction, functional container, etc.).

1. the language considered can be English or French,
2. the object must be a located object,
3. there is no condition involving the located object[5],
4. the highlighted property of the object must be that it plays the role of a located (physical) object in the current situation,[6]
5. there is no extra constraint.

We now summarize all the information gathered above. The information of this subsection tells us that the instantiation of the point conceptualization (recall that section 4.2.1 describes only point conceptualizations)[7] of any object occurs if the sentence in which this object is referred to is uttered in English or French, if it is a located object, and if the highlighted property of this object in the sentence considered — that is, the most relevant property of the object in this sentence — is that the object plays the role of a located object in the situation considered. Condition 3 and 5 are empty, which means that no condition involving the located object and no extra constraints need to be checked.

4.2.1.2 Object is a linear object which is focused in on one point

The provisional part of this case specifies that:

The object must be shaped as a line.

In the execution part, the conditions of the conceptual representation are the following:

1. the language considered can be English but not French,
2. the object must be a reference object,
3. the located object must be fixed and of size between 4 and 8,[8]
4. the highlighted property of the reference object must be that it is a linear object focused in on one point,
5. the extra constraint is that the speaker must be or think s/he is on a trajectory intersecting the reference object at the point focused in.

---

[5]The condition involving the located object applies only to the conceptual representations of reference objects. It is, therefore, obvious that no condition involving the located object is needed in the current case.

[6]Our system deals only with physical objects, geometrical objects (points, lines etc.), and accidents (holes, bends, breaks, etc.), all referred to as *Physical objects*. Events (e.g., a meeting), activities (e.g., the referent of *we were sleeping*), and states (e.g., *it was hot*) are not considered in our system. (These subdivisions and examples are all borrowed from Herskovits (1986) p. 127.)

[7]The other conceptualizations (line, surface, etc.) will be discussed in sections 4.2.2, 4.2.3, etc.

[8]Recall from chapter 3 that 4 is the approximate size of a prototypical tree; 5, the approximate size of a prototypical house; 6, the approximate size of a prototypical lake; 7, the approximate size of a seaside; and 8, the approximate size of a city.

Again, we summarize the information gathered above. This information tells us that the instantiation of the point conceptualization of an object shaped as a line occurs if the sentence in which it is refererred to is uttered in English, but not in French, if the object considered plays the role of a reference object, if the located object of the sentence considered (*not* the object considered, but the other object involved in the spatial relation) has the properties of being fixed and of being of approximate size greater than that of a human being, but smaller than or equal to that of a city, if the highlighted property of the reference object (the object considered) is that it is a linear object focused in on one point — in reality, the object is shaped as a line, but in the speaker's mind, only one point of this line is taken into consideration, and in his/her conceptual system, the entire object is this only point —, and if the speaker is or think s/he is going to intersect the reference object (the object considered) at the point s/he focused on. Recall that the knowledge needed to check Condition 5 is never available in our system, and the condition always assumed to be true.[9]

An example of a sentence where the reference object is conceptualized as a point, because it is a linear object focused in on one point is:

(59) The gas station is *at* the freeway. [Herskovits (1986) p. 138]
La station service est **à/à la jonction avec* l'autoroute.

Note that French, in this case, uses a different conceptualization: intersection (refer to case 4.2.12.1 for the French corresponding conceptual representation). In the example above, the different conceptualization used in French is illustrated by the fact that French uses the prepositional phrase 'à la jonction avec' (at the junction with), rather than 'à' alone.

We now explain the effect of each provision and condition listed at the beginning of the subsection.

Provision 1 of the provisional part eliminates some invalid cases. For example,

(60) * There is a gas station *at* the building.

is ruled out because a building is not shaped as a line, and therefore cannot be conceptualized as a point on a line.

In the execution part, Condition 3 eliminates additional invalid cases. For example,

(61) * The man is *at* the freeway.

is ruled out because a man is mobile,

(62) * The flower is *at* the freeway.

is ruled out because a flower is too small, and

[9] In subsequent sections, we will explain only the highlighted condition. As well, we will not need to indicate the role of the object (condition 2), as it will be uniformly that of a reference object.

(63) * The continent is *at* the freeway.

is ruled out because a continent is too large.

Finally, we compare this case with these of other researches: This conceptualization is the one assumed in sentences of the type described by Herskovits's case "Physical object on a line an indexically defined crosspath".

4.2.1.3 Object is a landmark in a highlighted medium

In the provisional part of the case, the following properties must be verified:

1. the object must be stationary or non-self-moving,
2. the noun the object is called must be preceded by an article,
3. the size of the object must be greater than 3.

In the execution part, the conditions of the conceptual representation are the following:

1. the language considered can be English or French,
2. the located object must be mobile
3. the highlighted property of the reference object must be that it is a landmark in a highlighted medium,
4. the extra constraint is that the located object must be on a defined trajectory intersecting the reference object.

The highlighted property (condition 3 of the execution case) says that among all the properties of the object considered, the one that is relevant is that this object is salient in a certain environment and therefore can be used to locate other objects of that environment. (Note that a large number of objects have a conceptual representation of this type.)

An example of a sentence where the reference object is conceptualized as a point because it is a highlighted medium is, in the context of a game:

(64) This time Jill will be *at* the tree, Paul *at* the bush, and Lew *at* the bench. [Herskovits (1986) p. 138]
Cette fois Jill sera *à* l'arbre, Paul *au* buisson, et Lew *au* banc.

In the provisional part, Provision 1 rules out invalid sentences of the type,

(65) * Jill is *at* the dog.

because a dog is self-moving and, for this reason, cannot be used to locate another object. Provision 2 eliminates sentences of the type

(66) Jill is *at* school.

which represent another scene and are taken care of by the case described in section 4.2.1.8. Note that the parser naturally rules out sentences in which the reference object is not preceded by an article as ungrammatical, unless the lexicon says that a given word might not always be preceded by an article. In the present case, no rule like provision 2 would have been needed, if words like 'school' or 'sea' could not appear without an article and have another meaning in these cases. Note further that Provision 2 was needed only because of English, since in French, all common nouns are preceded by an article, unless there is a morphological rule contradicting this fact, such as the morphological rule accompanying 'au' or 'en'. Provision 3 rules out

(67) * Jill is *at* the flower.

because a flower is too small relative to a girl.

Finally, in the execution part, Condition 2 rules out additional types of invalid sentences such as

(68) * The tree is *at* the house.

because a tree is immobile.

Finally, recall that like in all the cases where there is an extra constraint, this extra constraint is assumed to be verified by the system.

This conceptualization is the one assumed in sentences of the type described by Herskovits's case "Spatial entity at landmark in highlighted medium". Given our syntactic restrictions, however, our case involves only part of the one described by Herskovits.

4.2.1.4 Object is an exact location [This case is not implemented in the system]

In the provisional part, the following provision must be verified:

The object must have the natural property of being an exact location.

In the execution part, the conditions of the conceptual representation are the following:

1. the language considered can be English or French,
2. the located object must be bounded,
3. the highlighted property of the reference object must be that it is an exact location,
4. there is no extra constraint.

Condition 3 says that among all the properties of the object, the most relevant here is that it is an exact location.

An example of a sentences where the reference object is conceptualized as a point because it is an exact location is

(69) The book is *at* the place where you left it. [Herskovits (1986) p. 128]
Le livre est *à* l'endroit où tu l'as laissé.

In the provisional part, the unique Provision rules out sentences which describe a different scene and are taken care of by the case of section 4.2.1.5.

(70) The boy is *at* the supermarket.

is one such example.

In the execution part, Condition 2 rules out invalid sentences such as,

(71) * The sugar is *at* the place where you left it[10]

This conceptualization is the one assumed in some of the sentences of the type described by Herskovits's case "Spatial entity at location". Note, however that this case is more carefully restrictive than Herskovits's, since we are including only exact locations such as the ones specified by the words: 'place', 'spots', 'intersection', etc. In her case, Herskovits includes buildings, parks, etc. As mentioned in example (70), the type of sentences involving these entities are taken into consideration by us in section 4.2.1.5. This was done so because we believe that in these cases, 'at' has a functional value in addition to its locative value, which Herskovits did not consider.[11]

4.2.1.5 Object is a location with a purpose

In the provisional part, the following properties must be verified:

1. the object must have a social function,
2. the object must be of size greater than 3.[12]

In the execution part, the conditions of the conceptual representation are the following:

1. the language considered can be English or French,
2. the located object must be able to serve the function expected to be served at the reference object.
3. the highlighted property of the reference object must be that it is a location with a given purpose,
4. there is no extra constraint.

[10]The sentence sounds acceptable because 'the sugar' is understood as 'the container of sugar'. Our system, however, cannot recognize these cases of metonymy and rejects such sentences.

[11]As well, it might have been better to declare two ideal meanings for the preposition 'at': one purely geometrical, and the second one, mainly functional. We did not implement it, however.

[12]Recall that sizes were discussed in chapter 3.

Condition 3 says that the relevant property of the object considered here is that it must be a place (of size at least that of a building) in which certain given types of activities can be performed.

Examples of sentences where the reference object is conceptualized as a point because it is a location with a given purpose are:

(72) The boy *at* the supermarket.
Le garçon est *au* supermarché.

(73) She owns a cabin *at* the beach.
Elle a une cabane *à* la mer[13]

In these sentences, the boy has a reason to be at the supermarket and 'she' to own a cabin at the beach: the speaker assumes that the boy is probably shopping at the supermarket and 'she' usually spends her vacation in her cabin at the beach.

In the provisional part, Provision 1 rules out cases taken care of by the case of section 4.2.1.3 and 4.2.1.4 already discussed. For example, sentences like

(74) The boy is *at* the tree.

(75) The boy is *at* his spot.

are ruled out by Provision 1. Provision 2 rules out cases taken care of by the case of section 4.2.1.6 (to be discussed). For example,

(76) The boy is *at* the typewriter.

is ruled out by Provision 2, because a typewriter is of size smaller than 3.

In the execution part, Condition 2 rules out invalid sentences such as

(77) * The pigeon is *at* the supermarket.[14]

because, in a normal context, a pigeon cannot shop.[15]

This conceptualization is the one assumed in sentences of the type described by Herskovits's cases: "Spatial entity at location". Additionally, our case includes "Spatial entity at generic place". We organized these cases differently from Herskovits, because we believe that the social function of objects is taken into consideration in many of the cases where 'at' is used. Herskovits considers this point, only in her restrictive case, "Person using artifact".

---

[13]Note that 'mer' means 'sea' rather than 'beach'. However the French generic expression for 'at the beach' is 'à la mer'.

[14]Note that we did not use 'the dog' as a located object, because

(78) The dog is *at* the supermarket.

sounds acceptable. This might be because in western countries, a dog could be accompanying his/her master to the supermarket, which can be seen as serving the function expected by a dog at a supermarket, thereby satisfying condition 2. Our system, however, does not accept such sentences.

[15]Recall that the system assumes normality everywhere. See also the discussion of chapter 5 and 7 about whether there should be a distinction between normality and other types of conditions.

4.2.1.6 Object is an artifact with a given purpose

In the provisional part, the following properties must be verified:

1. the object must have a social function,
2. the object must be of size 1, 2, or 3,
3. the object must not have an inside.

In the execution part, the conditions of the conceptual representation are the following:

1. the language considered can be English or French,
2. the located object must be able to serve the function expected to be served at the reference object but is not supported by it.
3. the highlighted property of the reference object must be that it is an artifact with a given purpose,
4. there is no extra constraint.

Condition 3 says that the relevant property of the object considered here is that it must be an object (of size smaller than a house) which can be used to perform a certain type of activities.

Examples of sentences where the reference object is conceptualized as a point because it is an artifact with a given purpose are

(79) Maggie is *at* her desk. [Herskovits (1986) p. 135]
Maggie est *à* son bureau.

(80) The man is *at* the door.
L'homme est *à* la porte.

In these sentences, the located object is using the reference object in its normal use: Maggie is working at her desk and the man is attempting to enter through the door.

In the provisional part, Provision 1 acts as previously described in section 4.2.1.5, Provision 2 ensures that this case does not treat cases taken care of by case 4.2.1.5, Provision 3 rules out invalid sentences, such as

(81) * The man is *at* the bed.

because a bed has an inside.[16]

In the execution part, Condition 2 acts as previously described in section 4.2.1.5.

This conceptualization is the one assumed in sentences of the type described in Herskovits's case: "Person using artifact".

[16]Note that a bed does not *really* have an inside. What is considered its 'inside' is the space in between the mattress and the covers.

4.2.1.7 Object is a generic sea

In the provisional part, the following properties must be verified:

1. the object must be described by the word 'sea',
2. the name of the object must not be preceded by an article.

In the execution part, the conditions of the conceptual representation are the following:

1. the language considered can be English, but not French,
2. the located object must be smaller than 5,
3. the highlighted property of the reference object must be that it is a generic sea,
4. there is no extra constraint.

Condition 3 says that the relevant property in this case is that the reference object is a generic sea — by 'generic sea', we mean 'not a particular sea', but 'any sea'.

An example of a sentence where the reference object is conceptualized as a point because it is a generic sea is

(82) The marijuana containers are already *at* sea. [Herskovits (1986) p. 133]
Les conteneurs de marijuana sont déjà **à/en* mer.

Note that in this case, French uses a different conceptualization: functional container (refer to case 4.2.10.1 for the French corresponding conceptual representation).

In the provisional part, Provision 1 simply characterizes the case and Provision 2 rules out English and French sentences, such as

(83) The man is *at* the sea.

taken care of by case 4.2.1.3. In the execution part, Condition 2 makes sure that meaningless sentences, such as

(84) * The mountain is *at* sea.

are ruled out.

This conceptualization is the one assumed in sentences of the type described by Herskovits's case "Entity at sea".

4.2.1.8 Object is an institution one can be affiliated with and located inside it

In the provisional part, the following provision must be verified:

The object must be an institution.

In the execution part, the conditions of the conceptual representation are the following:

1. the language considered can be English or French,
2. the located object can serve one of the functions expected to be served at the reference object,
3. the highlighted property of the reference object must be that it is an institution one can be affiliated with and located inside it,
4. the extra constraint is that the located object has to be physically located at the reference object.

Condition 3 is self-explanatory, in this case.

Examples of a sentences where the reference object is conceptualized as a point because it is an institution one can be affiliated with and located inside are

(85) It is 3:00 pm; my son must still be *at* the University[17].
Il est 3 heures de l'après-midi; mon fils devrait encore être *à* l'université.

Recall that the system assumes that the extra constraint — the speaker thinks that his son is located at the university — is verified.

However, in the sentence

(86) My son studies *at* the University of Toronto.
Mon fils etudie *à* l'université de Toronto.

the located object does not have to be located at the university. This might be because 'the University of Toronto' is a proper name. Herskovits did not make a distinction at this point. We think it should be made, but did not implement it.

Note that in English the same idea can be expressed with the preposition 'in'. There seems to be a slight difference between 'at' and 'in' in this context, which is that 'at' sounds more natural when the located object is physically located at the reference object, and 'in' sounds more natural when the affiliation alone is referred to. French uses 'à/au' in every case.

The unique provision simply characterizes the case.

In the execution part, Condition 2 acts as previously described in section 4.2.1.5.

This conceptualization is the one assumed in sentences of the type described by Herskovits's case "Person at institution".

[17] 'School' seems to be an exception which must not be preceded by an article in the present case.

4.2.1.9 Horizontal non-low object involved in a passive support relation

In the provisional part, the following properties must be verified:

1. the object must be horizontal and non-low,
2. the object must have closed low or high sides.

In the execution part, the conditions of the conceptual representation are the following:

1. the language considered can be French, but not English,
2. the located object is non-adherent and cannot participate in an adherence relation with an object,
3. the highlighted property of the reference object must be that it is involved in a passive support relation,
4. there is no extra constraint.

Condition 3 says that the relevant property here is that the reference object must support another object, but that the support relation is not caused by any of the two objects. Instead, it is caused by a foreign body.

An example of a sentence where the reference object is conceptualized as a point because it is a horizontal and non-low object involved in a passive support relation is

(87) Le lustre est *au* plafond.
The lamp is **at/on* the ceiling.[18]

In this example, the support relation is passive because neither the lamp nor the ceiling are actively involved in the relation: the only active agents are the screws or the nails that hold the lamp to the ceiling, and these agents belong to neither the lamp, nor the ceiling. Note that in this case, English uses a different conceptualization: surface (refer to case 4.2.3.6 for the English corresponding conceptual representation).

In the provisional part, Provision 1 characterizes the case and eliminates the case taken care of by section 4.2.3.1, in which the English and the French objects are both conceptualized as a surface. Provision 2 rules out invalid sentences, such as

(88) * The poster is *on* the shade.

In the execution part, Condition 2 eliminates the English and French cases taken care of by section 4.2.3.5, in which the adherence relation is active or partly active. In English, this elimination has the effect of simply leaving this case to case 4.2.3.5; in French, since the conceptualization of case 4.2.3.5 is surface, whereas it is point here, the elimination has the effect of ruling out invalid French sentences in which the adherence relation is active and the conceptualization of the reference object used is a point. An example of such a sentence is

[18]Note that we show the French sentence first when the conceptualization is valid in French but not in English. We show the English sentence first, in the other cases.

(89) * La mouche est *au* mur [Vandeloise pp. 200-202].

This case was discussed by Vandeloise (1986 pp. 200-202) but not by Herskovits, because no distinction is made between an active and a passive support relation in English. The case discussed here is part of Herskovits's case "Spatial entity supported by physical object".

4.2.1.10 Object is a bounded surface and a country or a continent

In the provisional part, the following properties must be verified:

1. the object must be a country or a continent,

2. the object must be called by a proper name,

3. the name of the object must be masculine in French.[19]

In the execution part, the conditions of the conceptual representation are the following:

1. the language considered can be French, but not English,

2. there is no condition involving the located object,

[19]Note that in addition to the provisional part, in order for a conceptual representation to built, in general, some normality conditions must be satisfied (chapter 5). These normality conditions involve a size check, which checks that the size of the reference object is greater (or equal in certain cases) than the size of the located object, and a materiality check which makes sure that the located object is material. In the case of this section, however, the condition on materiality does not need to be checked, since a sentence such as

(90) There is a lot of wind *in* Canada

is acceptable. In most of the cases, the two conditions need to be checked, but in these, where one or the two conditions do not need to be checked, a footnote indicates it.

3. there are two[20] highlighted properties of the reference object: it must be a bounded surface[21], and a country or a continent[22],

4. there is no extra constraint.

Condition 3 says that the relevant properties of the reference object in this case are that it is a bounded surface — we assigned this type of general shape to objects with definite borders (see explanation at the beginning of section 4.2.4) — and that it is a country or a continent.

An example of a sentence where the reference object is conceptualized as a point because it is a bounded surface and a country or a continent is

(93) Le garçon est *au* Canada.
The boy is **at/in* Canada.

Note that in this case, English uses a different conceptualization: empty volume (refer to case 4.2.4.1 for the English corresponding conceptual representation).

In the provisional part, Provision 1 characterizes the case and eliminates in particular the cases treated by sections 4.2.1.11 in which French treats the case of cities and the part of 4.2.4.2, in which French treats geographic entities not referred to by their proper name. Provision 2 eliminates in particular the part of the cases of section 4.2.4.2 (in which geographical entities are referred to by their proper names). Provision 3 eliminates the French case of section 4.2.10.3, in which French treats the case of feminine

---

[20]When one case has two highlighted properties, its conceptualization can be matched to another one, even if only one of the highlighted properties is matched. This feature was implemented in an attempt to group together similar English cases and similar French ones. This attempt, however, was not carried out in every possible case (see section 4.2.14). In the cases of countries, continents, and cities, it was carried out: English conceptualizes in the same way all countries, cities, continents (provided that the continents are referred to by their proper names), and other entities. This conceptualization is described in cases 4.2.4.1 and 4.2.4.2. The highlighted property in this case is 'bounded surface'. In the case of French, a distinction must be made between countries, cities, and the other objects. For this reason, each different type of object is treated by a different case (sections 4.2.1.10, 4.2.1.11, 4.2.4.2, and 4.2.10.3). Because, we wanted to regroup all the English cases together, each French conceptual representation of these cases has two highlighted propérties: one that the English corresponding conceptual representation does not have: 'country or continent' or 'city', and a second one: 'bounded surface', which allows the matching to be performed. Note that the first highlighted property is not needed by the system, but was implemented for theoretical reasons: in order to distinguish between sentences of the type

(91) The boy is in Paris

and

(92) The boy is in Canada

which represent two different types of scenes.

[21]Note that bounded surface is both a property and a conceptualization. This might be confusing, but is nevertheless acceptable since a conceptualization is a schematization of the object, in which one of its properties is considered as if it was its only one. The same case of duality happens with other terms as well, such as 'edge of water body', 'edge of solid', 'full volume', etc.

[22]Note that, as mentioned in footnote no. 1, a Provision and a highlighted property might overlap.

countries and continents.[23] Note that French uses different conceptualizations in each of these four cases.

This conceptualization is the one assumed in some of the sentences of the type described by Herskovits as: "Spatial entity in area". Note that we divided her cases into four due to French convention with respect to names of geographical entities. These four cases are: a country or continent whose names are masculine in French and begin with a consonant (the current case study)[24], other countries or continents, cities, and all other bounded surfaces.

4.2.1.11 Object is a bounded surface and a city

In the provisional part, the following properties must be verified:

1. the object must be a city,
2. the object must be called by a proper name.

In the execution part, the conditions of the conceptual representation are the following:

1. the language considered can be French, but not English,
2. there is no condition involving the located object,
3. there are two highlighted properties of the reference object: it must be a bounded surface, and a city,
4. there is no extra constraint.

Condition 3 is self-explanatory, in this case. (See section 4.2.1.10 for the explanation of 'bounded surface', in this case.)

An example of a sentence where the reference object is conceptualized as a point because it is a city is

(94) Le garçon est *à* Paris.
The boy is **at/in* Paris.

Note that in this case, English uses a different conceptualization: empty volume (refer to case 4.2.4.1 for the corresponding English conceptual representation).

In the provisional part, Provision 1 and 2 eliminate all other cases in a way similar to that discussed previously in section 4.2.1.10.

This conceptualization is the one assumed in some of the sentences of the type described by Herskovits as: "Spatial entity in area" (see the discussion of the division of "Spatial entity in area" in section 4.2.1.10).

---

[23] Note that although we treat the cases of continents, countries, and cities as conceptual, we do not think that they should be handled like other cases because they seem purely conventional, rather than cognitive. Future work concerning this issue is discussed in chapter 7.

[24] Actually, the part about the starting letter was not implemented.

4.2.1.12 Object is an institution one can be affiliated with and whose location is irrelevant

In the provisional part, the following property must be verified:

The object must be an institution

In the execution part, the conditions of the conceptual representation are the following:

1. the language considered can be French, but not English,
2. the located object must be able to serve the function expected to be served at the reference object,
3. the highlighted property of the reference object must be that it is a type of institution one can belong to,
4. the location of the located object is irrelevant.

Condition 3 is self-explanatory in this case.

An example of a sentence where the reference object is conceptualized as a point because it is an institution one can be affiliated with is

(95) Joshua a deux ans, il n'est pas encore *à* l'école!
Joshua is two, he is not **at/in* school yet!

In this sentence, the location of Joshua is irrelevant: he could as well happen to be located in the school when the sentence is uttered.

Note that in this case, English uses a different conceptualization: empty volume. Note further that the English case requires that the object not be preceded by an article (refer to case 4.2.6.5 for the corresponding English conceptual representation).

In the provisional part, the unique Provision simply characterizes the case.

In the execution part, Condition 2 acts as mentioned previously in section 4.2.1.5, ruling out sentences of the type

(96) * The dog is in school.

This conceptualization is the one assumed in sentences of the type described by Herskovits as "Participant in an institution". Note that this case and case 4.2.1.8 are very close to one another. In fact, the only difference between them is the extra-sentential constraint. Since our system does not take this constraint into account, the scenes of section 4.2.1.8 and of this section are always built simultaneously (remember that in English, they are conceptualized differently in each case).

4.2.1.13 Object is a part of space or an environment

In the provisional part, the following provisions must be verified:

1. the object must have the natural function of creating an environment[25],
2. the object must not have the natural function of hiding,
3. the object must not have an almost nonexistent width.

In the execution part, the conditions of the conceptual representation are the following:

1. the language considered can be French, but not English,
2. there is no condition involving the located object,
3. there are two highlighted properties of the reference object: it must be a part of space or environment, and cannot hide,
4. there is no extra constraint.

Condition 3 says that the relevant properties of the reference object here are that it is a part of the space or an environment — such objects can be the sun, the wind, the corner of the room, etc. and that nothing can be hidden in it (unlike objects like the fog, or the dark).

An example of a sentence where the reference object is conceptualized as a point because it is a part of space or environment is

(97) Le garçon est *à* l'ombre.
The boy is **at/in* the shade.

Note that in this case, English uses a different conceptualization: empty volume (refer to case 4.2.6.6 for the English corresponding conceptual representation).

In the provisional part, Provision 1 simply characterizes the case. Provision 2 eliminates the case of section 4.2.6.7 (to be described) which treats sentences such as

(98) The boy is *in* the fog.
Le garçon est *dans le/*au* brouillard.

where the boy is hidden by the fog. This type of sentence must be ruled out, because French uses a different conceptualization in this case (see section 4.2.6.7). Provision 3 eliminates certain cases of section 4.2.6.7 (the cases treating the sentences corresponding to section 4.2.3.10 in French), such as

(99) The boy is *in* the picture.
Le garçon est **dans/sur* la photo.

[25] A few objects have this function: atmospherical objects (sun, shade, etc.), parts of space (the middle, the corner, etc.), and representations (the picture, the poster, etc.).

This type of sentences must be ruled out here, because French uses a different conceptualization in their case too.

This conceptualization is the one assumed in part of the sentences of the type described by Herskovits as "spatial entity in part of space or environment". Note that in her case, Herskovits does not differentiate between an environment in which one can hide and one in which one cannot. This is because the distinction is not made in English.

4.2.2 The conceptualizations as a line

In each of the conceptual representations described in this section, the node pointed to by the arc META contains the information specifying that the dimension of the metaphor is LINE, and its fullness and width are NIL. This conceptualization is geometric and indicates that the object considered is conceptualized as a line.

Recall that all the conceptual representations discussed here are these of reference objects, since the only conceptual representation of a located object was discussed in section 4.2.1.

4.2.2.1 Object is a supporting line

In the provisional part, the following properties must be verified:

1. the object must be shaped as a line,
2. the object must have the natural function of supporting.

In the execution part, the conditions of the conceptual representation are the following:

1. the language considered can be English or French,
2. the located object must be light enough to be supported by the reference object,
3. the highlighted property of the reference object must be that it is a linear support,
4. there is no extra constraint.

Condition 3 says that the relevant properties of the reference object are that it has the general shape of a line and the natural function of supporting.

An example of a sentence where the reference object is conceptualized as a line because it is a supporting line is

(100) The shirt is *on* the washing line.
La chemise est *sur* le fil à linge.

In the provisional part, Provisions 1 and 2 characterize the case. Provision 1 rules out in particular the case treated by section 4.2.3.1, in which the object is a surface. Provision 2 rules out in particular the cases treated in sections 4.2.2.2, 4.2.2.3, and 4.2.2.4, in which

no positive support is involved: (mainly) contiguity.

In the execution part, Condition 2 rules out invalid sentences, such as

(101) * The elephant is *on* the washing line.

This conceptualization is the one assumed in some of the sentences of the type described by Herskovits's case: "Spatial entity supported by physical object". We differ from Herskovits in that she assumes that support cannot be achieved by an object conceptualized as a line. When an object shaped as a line is a support, she claims that it is in fact conceptualized as a (thin) surface [Herskovits (1986) p.141]. We do not maintain this claim.

4.2.2.2 Object is a fixed line of reference

In the provisional part, the following properties must be verified:

1. the object must be shaped as a line,
2. the object must not be a body of water,
3. the object must be stationary.

In the execution part, the conditions of the conceptual representation are the following:

1. the language considered can be be English or French,
2. the located object must be fixed and of size greater than 3,
3. the highlighted property of the reference object must be that it is a fixed line of reference,
4. there is no extra constraint.

Condition 3 says that the relevant property of the reference object in this case is that it has the general shape of a line and allows other objects to be contiguous to it (no support relation is involved here).

An example of a sentence where the reference object is conceptualized as a line because it is a fixed line of reference is

(102) A village *on* the border ...
Un village *sur* la frontière ...

In the provisional part, Provisions 1 eliminates the cases taken care of in sections 4.2.3.2 in which the object is a surface. Provision 2 eliminates the case of section 4.2.2.3 and 4.2.2.4, in which special types of objects are involved (edges of water and edges of other geographical entities, respectively). This distinction is made because French uses different types of conceptualization in these cases. Provision 3 rules out invalid sentences, such as

(103) * The house is *on* the snake.[26]

In the execution part, Condition 2 rules out invalid sentences of the type

(104) * The truck is *on* the border.

or

(105) * The flower is *on* the border.

This conceptualization is the one assumed in sentences of the type described by Herskovits's case "Physical or geometrical object contiguous with a line".

4.2.2.3 Object is a body of water whose edge is focused on

In the provisional part, the following properties must be verified:

1. the object must have a bounded low or high outside surface,
2. the object must be a body of water.

In the execution part, the conditions of the conceptual representation are the following:

1. the language considered can be English but not French,
2. the located object must be fixed and of size between 4 and 8,
3. the highlighted property of the reference object must be that it is a body of water whose edge is focused on,
4. there is no extra constraint.

Condition 3 says that the relevant property of the reference object in this case is that while the body of water is entirely referred to, in the conceptual system of the speaker, only the edge is represented. Note that all bodies of water are considered to have an edge.

An example of a sentence where the reference object is conceptualized as a line because it is a body of water whose edge is focused on, is

(106) The house *on* the lake ...
La maison **sur le/au bord du* lac ...

Note that, in this case, French uses a different conceptualization: water-body-edge (refer to case 4.2.8.1 for the corresponding French conceptual representation).

In the provisional part, Provision 1 characterizes the case and eliminates the case of section 4.2.2.1, in which the object is only linear: a body of water is never only linear[27],

[26] There are two problems with this example: the snake is mobile and the snake is too small to be a reference for a house. It is the first problem we attempted to demonstrate here.

[27] A river, for example, is both linear (when considered in its integrity) and planar (when considered at a given part of it). It can therefore be conceptualized by the case of this section *and* the case of section 4.2.2.1.

and Provision 2 eliminates the cases of section 4.2.2.4, in which the object is not a body of water.

In the execution part, Condition 2 acts as previously described in section 4.2.2.2.

This conceptualization is the one assumed in part of the sentences of the type discussed by Herskovits's case "Physical object contiguous with edge of geographical area". Mark (1990), however, discusses this case and replaces it by "Figure on edge-of(Ground)", together with a list of exceptional cases which do not need the edge name to be specified (Mark (1990), p. 7). Mark also distinguishes the case of edges of water bodies and edges of solid geographic entities in French (case 4.2.2.4 as well). We adopt this distinction as well. Mark observes that in this case, there is a slight difference between European French and Quebecois French, where in the former, the edge noun might not be obligatory, whereas it seems unusual to omit it in the latter (Mark 1990, p. 10).

4.2.2.4 Object is a solid geographic entity whose edge is focused on

In the provisional part, the following properties must be verified:

1. the object must be an open space,
2. the object must not be a body of water,
3. the object must be of size smaller than 9.

In the execution part, the conditions of the conceptual representation are the following:

1. the language considered can be English but not French,
2. the located object must be fixed and of size between 4 and 8.
3. the highlighted property of the reference object must be that it is an open space whose edge is focused on,
4. there is no extra constraint.

Condition 3 says that the relevant property of the reference object in this case is that it is an open space whose edge is focused on. Note that here, unlike in case 4.2.2.3, not every open-space has an edge. For example, a city does not have a well enough defined edge.[28]

An example of a sentence where the reference object is conceptualized as a line because it is a solid geographic entity whose edge is focused on is

(107) The house *on* the park ...
La maison **sur le /en bordure du* parc ...

[28] This seems to contradict the place in section 4.2.1.10 where we said that countries, continents, and cities are well defined by their borders, but it does not: borders might not be solid, whereas edges are.

Note that in this case, the French conceptualization is different: open-space-edge (refer to case 4.2.9.1 for the corresponding French conceptual representation).

In the provisional part, Provision 1 eliminates the cases of sections 4.2.2.2 in which the object is shaped as a line, and Provision 2 eliminates the case of section 4.2.2.3, in which bodies of water are considered. Provision 3 eliminates the case of entities that are too large to be conceptualized in the way described here.

In the execution part, Condition 2 acts as was previously mentioned in section 4.2.2.2.

This conceptualization is the one assumed in part of the sentences of the type discussed in Mark's study as "Figure on edge-of(Ground)" (see case 4.2.2.3). (The same remark as in case 4.2.2.3 about Quebecois French applies to this case.)

4.2.3 The conceptualizations as a surface

In each of the conceptual representations described in this section, the node pointed to by the arc META contains the information specifying that the dimension of the metaphor is SURFACE, its fullness and its width are NIL. This conceptualization is geometric and indicates that the object considered is conceptualized as a surface. There is a difference between SURFACE and BOUNDED SURFACE, which will be discussed at the beginning of section 4.2.4.

Recall that all the conceptual representations discussed here are these of reference objects, since the only conceptual representation of a located object was discussed in section 4.2.1.

4.2.3.1 Object is a plane horizontal support

In the provisional part, the following properties must be verified:

1. the object must be a horizontal surface,
2. the object must not be a geographic object (but can be a roadway),
3. the object must not be a geometric object,
4. the object must not be a page division.

In the execution part, the conditions of the conceptual representation are the following:

1. the language considered can be English and French,
2. the located object must be light enough to be supported by the reference object,
3. the highlighted property of the reference object must be that it is a plane horizontal support,
4. there is no extra constraint.

Condition 3 says that the relevant property of the reference object in this case is that it has a plane horizontal part which can support other objects.

An example of a sentence where the reference object is conceptualized as a surface because it is a plane horizontal support is

(108) The glass is *on* the table.
Le verre est *sur* la table.

In the provisional part, Provision 1 characterizes the case and, in particular, eliminates the cases treated by sections 4.2.1.9, 4.2.3.5, and 4.2.3.6, in which the object has a horizontal shape and is not low (i.e., a vertical or a high-horizontal one, such as that of a ceiling), rather than an horizontal one. Provision 2 eliminates the cases of sections 4.2.4.1, 4.2.1.10, and 4.2.1.11, in which the object is a continent, a country, or a city, and is conceptualized as a bounded surface, rather than a surface. Provisions 3 and 4 eliminate the cases of section 4.2.4.2, which are cases where the objects are also conceptualized as a bounded surface, rather than a surface. Note that Provisions 2, 3, and 4 eliminate the same type of cases, in which the objects are conceptualized as a bounded surface, rather than a surface. These objects are conceptualized differently because their definition depend on their limits (see beginning of section 4.2.4).

In the execution part, Condition 2 rules out sentences of the type

(109) * The elephant is *on* the flower.

This conceptualization is the one assumed in sentences of the type described by Herskovits's case "Spatial entity supported by physical object".

4.2.3.2 Object is a fixed surface of reference

In the provisional part, the following properties must be verified:

1. the object must have a surface,
2. the object must not be a geographic object,
3. the object must not be a geometric object,
4. the object must not be a page division,

In the execution part, the conditions of the conceptual representation are the following:

1. the language considered can be English and French,
2. the located object must have no weight or seem to have no weight,
3. the highlighted property of the reference object must be that it is a fixed surface of reference,
4. there is no extra constraint.

Condition 3 says that the relevant property of the reference object in this case is that it is planar and can have other objects contiguous with it. No support is involved in this case.

An example of a sentence where the reference object is conceptualized as a surface because it is a fixed surface of reference is

(110) Do not put your dirty fingers *on* my clean suit! [Herskovits (1986) p. 144]
Ne met pas tes doigt sales *sur* mon costume tout propre!

In the provisional part, Provisions 1, 2, 3, and 4 simply characterize the case.

In the execution part, Condition 3 eliminates the case where the reference object is an actual support, such as in

(111) The glass is on the table.

which are treated by section 4.2.3.1.

This conceptualization is the one assumed in sentences of the type described by Herskovits as "Physical object contiguous with another".

4.2.3.3 Object is a part of another, which it can support

In the provisional part, no provision needs to be verified.[29]

In the execution part, the conditions of the conceptual representation are the following:

1. the language considered can be English and French,
2. the reference object must be part of the located object,
3. the highlighted property of the reference object must be that it is a part of another object which it can support,
4. there is no extra constraint.

Condition 3 is self-explanatory, in this case.

An example of a sentence where the reference object is conceptualized as a surface because it is part of another, which it can support is

(112) A table *on* three legs ... [Herskovits (1986) p.146]
Une table *sur* trois pieds ...

[29] The normality condition involving the size of objects does not need to be checked here because the sizes do not matter (they should somehow, or at least their weight should, but our system is designed in a way where they do not), since a boy can be on his head, even though a head is smaller than the body of the boy. (This should be improved.)

In the execution part, Condition 2 rules out invalid sentences such as

(113) * The table *on* the cup ...

because a cup is not part of a table. This conceptualization is the one assumed in sentences of the type described as one of Herskovits's cases: "Physical object on part of itself".

4.2.3.4 Object is an uncovered geographic entity

In the provisional part, the following property must be verified:

The object must be a geographic entity[30].

In the execution part, the conditions of the conceptual representation are the following:

1. the language considered can be English or French,
2. the located object must be atmospheric,
3. the highlighted property of the reference object must be that it is an uncovered geographic entity,
4. there is no extra constraint.

Condition 3 says that the relevant property of the reference object in this case is that it is a geographic entity which could look like it is covered by another object. (Note that the last property is not restrictive: any geographical entity might look this way.)

An example of a sentence where the reference object is conceptualized as a surface because it is an uncovered geographic entity is

(114) The rain *on* Toronto ...
La pluie *sur* Toronto ...

In the provisional part, the unique Provision rules out invalid sentences, such as

(115) * The rain *on* the box ...

because a box is not a geographical entity. In the execution part, Condition 2 rules out invalid sentences, such as

(116) * The bird *on* Toronto ...

because a bird is not atmospheric. This conceptualization is the one assumed in sentences of the type described as one of Herskovits's cases: "Physical object over a geographic object".

[30]The normality condition on the materiality of the located object does not have to be satisfied here (on the contrary in several cases, the located object must be immaterial), because the located object needs to be atmospheric, and often, atmospheric entities are immaterial.

4.2.3.5 Horizontal non-low object is involved in an active support relation

In the provisional part, the following properties must be verified:

1. the object must have closed sides (high or low),
2. the object must not be horizontal.

In the execution part, the conditions of the conceptual representation are the following:

1. the language considered can be English or French,
2. the located object must be adherent or able to participate in an adherence relation,
3. the highlighted property of the reference object must be that it is a physical object involved in an active support relation,
4. there is no extra constraint.

Condition 3 says that the relevant property of the reference object in this case is that it can have another object adhere to it, and one of the two objects is responsible at least in part for this adherence relation.

An example of a sentence where the reference object is conceptualized as a surface because it is a horizontal non-low object involved in an active support relation is

(117) There is water *on* the wall.
Il y a de l'eau *sur* le mur.

In the provisional part, Provision 1 rules out invalid sentences such as

(118) * The poster is *on* the shadow.

and Provision 2 eliminates the case of section 4.2.3.1, in which the object is horizontal.

In the execution part, Condition 2 eliminates case 4.2.3.6 where the object is involved in a passive support relation, if the sentence is expressed in English and rules out invalid French sentences of the type

(119) * Le lustre est *sur* le plafond.[31]

This case is discussed by Vandeloise (1986 p.200-202). Herskovits does not discuss it, because in English, no distinction is made between active and passive support relations. Herskovits simply includes this case among others, in the category "Spatial entity supported by physical object".

[31] It sounds like Condition 2 has two different roles, but in fact, it has only one. The reason why it looks like it has several ones is that French uses a different conceptualization in the case of an active and a passive adherence relation, whereas English uses only one in both cases.

4.2.3.6 Horizontal non-low object is involved in a passive support relation

This section describes the English conceptual representation corresponding to the French one of section 4.2.1.9. Since the Provisional and Execution information is exactly the same in this case as in the case of section 4.2.1.9, it is not repeated here. In this section, we simply recall the example of section 4.2.1.9[32]

(120) The lamp is *on* the ceiling.
Le lustre est **sur/au* plafond.

and recall that the English conceptualization of this case is surface, whereas the French one is point.

4.2.3.7 Object is a horizontal support and has sides

In the provisional part, the following properties must be verified:

1. the object must have an inside surface,
2. the object must have sides,
3. the bottom surface of the object must be highlighted.

In the execution part, the conditions of the conceptual representation are the following:

1. the language considered can be English, but not French,
2. the located object must be light enough to be supported by the reference object,
3. there are two highlighted properties of the reference object: it must be a horizontal support and have sides,
4. there is no extra constraint.

Condition 3 says that the two relevant properties of the reference object in this case are that it has both a bottom surface which can support, and an inside (or merely sides).

Examples of sentences where the reference object is conceptualized as a surface because it is an horizontal support and has sides are

(121) Mom is *on* the bus.
Maman est **sur/dans* le bus.

(122) I live *on* a farm.
J'habite **sur/dans* une ferme.

(123) I was *on* the street.
J'étais **sur/dans* la rue.

[32]Note that this time we show the English sentence first, and the French one second, because the conceptual representation described here applies to English.

Note that in this case, French uses a different conceptualization: empty volume (refer to case 4.2.6.12, for the French corresponding conceptual representation). One of our observations is that in the case where an object has physical sides (could be liquid sides too), French uses the empty volume conceptualization. If, however, an object does not have sides, then French does not use this conceptualization (whereas, English does, as in "The boy is in the shade.")

In the provisional part, Provisions 1, 2, and 3 simply characterize the case.

In the conditional part, Condition 2 makes sure that a support relation can occur by ruling out sentences of the type

(124) * The elephant is *on* the bus.

This conceptualization is the one assumed in sentences of the type described by Herskovits's case "Physical object transported by a large vehicle", together with other sentences, not specifically discussed by Herskovits, but which we assume are included in "Spatial entity located on geographical location". We saw similarities between these various sentences and attempted to group them together.

4.2.3.8 Object is a wall, which can be contiguous with other objects

In the provisional part, the following properties must be verified:

1. the object must be a 'wall',
2. The object should be qualified with an adjective (Herskovits 1986 p. 145-146).

In the execution part, the conditions of the conceptual representation are the following:

1. the language considered can be English, but not French,
2. the located object must be fixed,
3. the highlighted property of the reference object must be that it is a wall which can be contiguous with other objects,
4. there is no extra constraint.

Condition 3 is self-explanatory, in this case.

An example of a sentence where the object is conceptualized as a surface because it is a wall which can be contiguous with other objects is

(125) *On* the left wall, there is a chest of drawers.[33] [Herskovits (1986) p. 145]
**Sur/Contre* le mur de gauche, se trouve une commode.

[33]Note that 'on' cannot be used in this way in Australian English: 'by' must be used instead.

Note that in this case, French uses a different conceptualization: opposing force (refer to case 4.2.13.1 for the corresponding French conceptual representation).

In the provisional part, Provisions 1 and 2 simply characterize the case.

In the execution part, Condition 2 rules out invalid sentences such as

(126) * The dog is *on* the left wall.

This conceptualization is the one assumed in sentences of the type described as one of Herskovits's cases: "Physical object contiguous with a wall".

4.2.3.9 Object is an open-space support

In the provisional part, the following properties must be verified:

1. the object must be an open space,
2. the object must have a low surface,
3. the object must not be called by a proper name,

In the execution part, the conditions of the conceptual representation are the following:

1. the language considered can be English or French,
2. there are no conditions on the located object,
3. the highlighted property of the reference object must be that it is an open space support,
4. there is no extra constraint.

Condition 3 is self-explanatory, in this case.

An example of a sentence where the reference object is conceptualized as a surface because it is an open-space support is

(127) My house is *on* the continent.
Ma maison est *sur* le continent.

In the provisional part, Provisions 1 and 2 simply characterize the cases (they might be redundant, though). Provision 3 eliminates the case of section 4.2.4.1, in which the object is referred to by its proper name, which causes the use of a different preposition, in the case of a continent, in both English and French.[34]

This conceptualization is the one assumed in sentences of the type described as one of Herskovits's cases: "Physical object located on an open space".

[34]The reason why a continent is conceptualized as a bounded surface when it is referred to by its name, and a surface otherwise, might be that in the first case, it is thought of as a political entity for which the borders are relevant, and in the second case, it is thought of as a geographical entity, so large that the borders do not matter: it is its size that is prominent. Another reason could also be that the borders of continents are usually natural, whereas the borders of countries might be artificial.

4.2.3.10 Object is an environment and has a tiny width

In the provisional part, the following properties must be verified:

1. the object must have the natural function of creating an environment,
2. the object has an almost non-existent width.

In the execution part, the conditions of the conceptual representation are the following:

1. the language considered must be French, but not English,
2. there is no condition involving the located object,
3. there are two highlighted properties of the reference object: it must be a part of space or environment, and have almost no width.
4. there is no extra constraint.

Condition 3 is self-explanatory in this case.

An example of a sentence where the reference object is conceptualized as a surface because it is an environment and has a tiny width is

(128) Le garçon est *sur* la photo.
The boy is **on/in* the picture.

Note that in this case, English uses a different conceptualization: empty volume with a tiny width (refer to case 4.2.7.1 for the corresponding English conceptual representation). (The English case, however, could be grouped together with the cases of sections 4.2.6.6 and 4.2.6.7.)

In the provisional part, Provisions 1 and 2 eliminate cases of sections 4.2.1.13, which takes care of French sentences such as

(129) Le garçon est *à* l'ombre.
The boy is **at/in* the shade.

and 4.2.6.8 which takes care of French sentences such as

(130) Le garçon est *dans* le brouillard.
The boy is *in* the fog.

Note that in English, the sentences

(131) The boy is *in* the fog.

(132) The boy is *in* the shade.

(133) The boy is *in* the picture.

should all belong to the same case (currently, however, they are treated by three different cases: the cases of section 4.2.6.6, 4.2.6.7, and 4.2.7.1), whereas in French, their translation belong to three different cases (4.2.6.8, 4.2.1.13, and 4.2.3.10, respectively).

This conceptualization is not described by Herskovits, who does not differentiate representations such as pictures from other objects. Herskovits instead includes the case of representations in the case "Accident/object part of physical or geometric object".

4.2.4 The conceptualizations as a bounded surface

In each of the conceptual representations described in this section, the node pointed to by the arc META contains the information specifying that the dimension of the metaphor is BOUNDED SURFACE, and its fullness and width are NIL. This conceptualization is geometrical but involves some functional property, as described below. It indicates that the object considered is conceptualized as a BOUNDED SURFACE. Note that a SURFACE and a BOUNDED SURFACE are different types of conceptualizations. Objects with definite boundaries of importance for their definition are conceptualized as BOUNDED SURFACES, rather than plain SURFACES. This is the case for countries, for example, where a country is defined in terms of its borders. It is the case in the domain of geometry as well, where, for example, a rectangle would not be a rectangle if it did not have well-defined sides. In domains where the borders are not relevant to the definition of the object, the conceptualization used is SURFACE. In the case of a table, for example, even though a table is a bounded object, since its boundaries are not what makes that object a table, a table is conceptualized as a SURFACE, rather than as a BOUNDED SURFACE.

Recall that all the conceptual representations discussed here are these of reference objects, since the only conceptual representation of a located object was discussed in section 4.2.1.

4.2.4.1 Object is a bounded surface and a country, a continent, or a city

In the provisional part, the following properties must be verified:

1. the object must be geographic,
2. the name of the object must be a proper noun.

In the execution part, the conditions of the conceptual representation are the following:

1. the language considered must be English, but not French,
2. there is no condition involving the located object,
3. there are two highlighted properties of the reference object: it must be a bounded surface, and a country or continent,
4. there is no extra constraint.

Condition 3 is self-explanatory in this case.

Examples of sentences where the reference object is conceptualized as a bounded surface because it is a bounded surface, a continent, a country, or a city, are

(134) The boy is *in* Venezuela.
Le garçon est **dans le/au* Vénézuéla.

(135) The boy is *in* Africa.
Le garçon est **dans l'/en* Afrique.

(136) The boy is *in* Paris.
Le garçon est **dans la/à* Paris.

(137) The boy is *in* Afghanistan.
Le garçon est **dans le/en* Afghanistan.

Note that in this case, French uses different conceptualizations: the object is conceptualized as a point or functional container. It is conceptualized as a functional container if its name is feminine or if it is masculine and starts with a vowel (refer to case 4.2.10.3), and it is point, otherwise (refer to cases 4.2.1.10 and 4.2.1.11).

In the provisional part, Provisions 1 and 2 eliminate the cases of section 4.2.4.2, which considers the case where the object is not referred to by its proper name and is geographic, geometric, or a page division (in this case, French uses the bounded surface conceptualization, rather than the functional container one).

This conceptualization is the one assumed in part of the sentences of the type described by Herskovits as "Spatial entity in area". However, as mentioned earlier, Herskovits does not differentiate between geographical areas and others because English does not make this distinction. Since French does, and uses different prepositions, we distinguish these cases (see section 4.2.1.10).

4.2.4.2 Object is a bounded surface

In the provisional part, the following properties must be verified:

1. the object must be geographic and social (country, city, etc. rather than sea, mountains, etc.), but not an open-space (continent) and not a proper name,
2. or the object must be geometric,
3. or the object must be a page division.

In the execution part, the conditions of the conceptual representation are the following:

1. the language considered must be English or French,
2. there is no condition involving the located object,

3. the highlighted property of the reference object must be that it be a bounded two-dimensional area,

4. there is no extra constraint.

Condition 3 in this case, is self explanatory.

Examples of sentences where the reference object is conceptualized as a bounded surface because it is a bounded surface and an object of the type considered here are

(138) The mark is *in* the margin.
Le signe est *dans* la marge.

(139) The girl is *in* the country
La fille est *dans* le pays.

Note that in this case, French conceptualizes 'pays' in the same way as English does 'country'. In the provisional part, Provision 1 eliminates the case of section 4.2.4.1, in which countries, continents, and cities referred to by their proper name are dealt with, and section 4.2.3.9, in which continents are dealt with. Again, this distinction is made because French uses different conceptualizations in both cases (see section 4.2.1.1).

This conceptualization is the one assumed in part of the sentences of the type described by Herskovits as "Spatial entity in area". As mentioned, in the previous case (4.2.4.1), we distinguish between geographic entities called by their proper name and the other types of entities, whereas Herskovits does not.

4.2.5 Conceptualization as a full volume

In each of the conceptual representations described in this section, the node pointed to by the arc META contains the information specifying that the dimension of the metaphor is VOLUME, its fullness is FULL and its width is NIL. This conceptualization is geometric and functional and indicates that the object considered is conceptualized as if it were a container and as if its constituents were objects located in it.

Recall that all the conceptual representations discussed here are these of reference objects, since the only conceptual representation of a located object was discussed in section 4.2.1.

4.2.5.1 Object is apprehended as a whole

In the provisional part, the following provision must be verified:

The object must be a full volume.[35]

In the execution part, the conditions of the conceptual representation are the following:

[35] Note that the normality condition concerning the materiality of the located object does not need to be satisfied here. This is in order to allow sentences such as

1. the language considered can be English or French,
2. the located object must be allowed to be considered as part of the reference object,
3. the highlighted property of the reference object must be that it is apprehended as a whole,
4. there is no extra constraint.

Condition 3 says that the relevant property of the reference object in this case is that it is an object which is seen as having other objects embedded in it (even though these other objects might be its constituents).

Examples of sentences where the reference object is conceptualized as a full volume because it is apprehended as a whole are

(141) The hole *in* the shirt ...
Le trou *dans* la chemise ...

(142) The nail *in* the board ...
Le clou *dans* la planche ...

In the provisional part, the unique Provision simply characterizes the case.

In the execution part, Condition 2 rules out invalid sentences such as

(143) * The cavity *in* the paper ...

because paper cannot have cavities, or

(144) * The fish *in* the air ...

because fish are not part of the constituents of the air. This conceptualization is the one assumed in some of the sentences of the type described by Herskovits as "Gap/Object embedded in physical object".

4.2.6 The conceptualizations as an empty volume

In each of the conceptual representations described in this section, the node pointed to by the arc META contains the information specifying that the dimension of the metaphor is VOLUME, its fullness is EMPTY, and its width is NIL. This conceptualization is geometrical and functional, and indicates that the object (together with its constituents) is considered as a container that can contain other objects foreign to it.

Recall that all the conceptual representations discussed here are these of reference objects, since the only conceptual representation of a located object was discussed in section 4.2.1.

---

(140) The hole *in* the shirt is big.

where a hole is immaterial. As well, the condition concerning the relative sizes of the objects does not matter because the constituents of objects are either immaterial or, in the case of holes, can be of any size.

4.2.6.1 Object is a well-defined low container

In the provisional part, the following properties must be verified:

1. the object must have an empty low inside,
2. the name of the object must be preceded by an article.[36]

In the execution part, the conditions of the conceptual representation are the following:

1. the language considered can be English or French,
2. the reference object must be hermetic to the located object,
3. the highlighted property of the reference object must be that it is a well-defined low container,
4. there is no extra constraint.

Condition 3 says that the relevant property of the reference object in this case is that it is an object which has a well-defined inside in its lower part, if it is composed of two parts, one above the other, and simply a well-defined inside, if it is made of only one part. This means that in the case of a tree, for example, where the tree can be divided into a low part (the trunk) and a high part (the branches), the part considered here is the trunk. In the case of a glass, which has only one part, the part considered here is the entire glass.

An example of a sentence where the reference object is conceptualized as an empty volume because it is a defined low container is

(145) The water is *in* the glass.
L'eau est *dans* le verre.

In the provisional part, Provision 1 eliminates cases 4.2.6.2, 4.2.6.3, and 4.2.6.4 for theoretical purposes.[37] These are cases in which the object considered is the higher part of an object and/or is not a well-bounded container. Provision 2 eliminates the cases of section 4.2.6.5 involving an institution, which has a different conceptualization in French when it corresponds to the case where it is not preceded by an article in English. For example, sentences of the type

(146) The boy is *in* school.
Le garçon est **dans/à* l'école.

[36] The normality condition concerning the materiality of the located object does not need to be satisfied here, because in certain cases the reference object can contain gas, which is not considered to be material.

[37] By this, we mean that there is no difference within English or French or between English and French which caused our distinction: the conceptualization is the same in the current case and the three following ones (4.2.6.2, 4.2.6.3, and 4.2.6.4). The reason why we did, however, distinguish each of the four cases from the others is that different scenes are represented every time, and the sentence has a different meaning in each case.

are eliminated here.

In the execution part, Condition 2 rules out invalid sentences, such as

(147) * The water is *in* the bag.

because a bag is not hermetic to water and therefore cannot contain water.[38] This conceptualization is the one assumed in some of the sentences of the type described by Herskovits as "Spatial entity in container". Herskovits, however does not differentiate between low and high insides. Distinguishing them allows us to differentiate, for example, the two meanings of "The boy is in the tree", where the boy is either in the trunk of the tree (lower part), or in its branches (higher part), which Herskovits does not.

4.2.6.2 Object is a not a well-defined low container

In the provisional part, the following properties must be verified:

1. the object must have an empty vague low inside,
2. the name of the object must be preceded by an article.

In the execution part, the conditions of the conceptual representation are the following:

1. the language considered must be English or French,
2. the reference object must be hermetic to the located object,
3. the highlighted property of the reference object must be that it is not a well-defined low container.
4. there is no extra constraint.

Condition 3 says that the relevant property of the reference object in this case is that it is an object which has an inside in its lower part, if it is composed of two parts, one above the other, and simply an inside, if it is made of only one part. In this case, however, the inside is different from the inside of case 4.2.6.1, in that it is not as well delimited: it is human beings who have the capacity of considering this part of the object an inside.

An example of a sentence where the reference object is conceptualized as an empty volume because it is not a well-defined low container is

(148) There is a worm *in* the strawberries.
Il y a un ver *dans* les fraises.

In this section, 'the strawberries' do not constitute a well-defined inside.

Provision 1 distinguishes this case from the case of section 4.2.6.1, where the object

[38] Note, however, that there could be water in a bag, by accident. In our system, however, the default mode is normal, and the accidental cases are not taken care of.

has the property of being a well-defined container. Provision 2 works as described previously in section 4.2.6.1.

Condition 2 works as described previously in section 4.2.6.1.

This conceptualization is the one assumed in some of the sentences of the type described by Herskovits as "Physical object in the outline of another, or of a group of objects". Herskovits, however, does not differentiate between low and high. (See the last remark of section 4.2.6.1 for relevant discussion.)

4.2.6.3 Object is a well-defined high container

In the provisional part, the following properties can be verified:

1. the object must have an empty high inside,
2. the name of the object must be preceded by an article.[39]

In the execution part, the conditions of the conceptual representation are the following:

1. the language considered must be English or French,
2. the reference object must be hermetic to the located object,
3. the highlighted property of the reference object must be that it is a well defined high container.
4. there is no extra constraint.

Condition 3 says that the relevant property of the reference object in this case is that it is an object which has a well-defined inside in its higher part, if it is composed of two parts, one above the other. This conceptual representation is not built if the object does not have two parts one above the other.

We could not find an example of this case. We included it, however, because it seems to us that examples could be found since some were found for the cases of sections 4.2.6.1, 4.2.6.2, and 4.2.6.4 which are similar to this case.

See section 4.2.6.1 for a description of the effect of the similar provisions and conditions.

This conceptualization is the one assumed in some of the sentences of the type described by Herskovits as "Spatial entity in container". Herskovits, however does not differentiate between low and high. (See the last remark of section 4.2.6.1 for a relevant discussion)

[39]The normality condition concerning the materiality of the located object does not need to be satisfied here, because in certain cases, the reference object can contain gas, which is not considered to be material.

4.2.6.4 Object is not a well-defined high container

In the provisional part, the following properties must be verified:

1. the object must have an empty vague high inside,
2. the name of the object must be preceded by an article.

In the execution part, the conditions of the conceptual representation are the following:

1. the language considered must be English or French,
2. the reference object must be hermetic to the located object,
3. the highlighted property of the reference object must be that it is not a well-defined high container.
4. there is no extra constraint.

Condition 3 says that the relevant property of the reference object in this case is that it is an object which has a well-defined inside in its higher part, if it is composed of two parts, one above the other. In this case, however, the inside is different from the inside of case 4.2.6.3, in that it is not as well delimited: it is the human beings who have the capacity of considering this part of the object an inside. Again, if the object is not constituted of two parts one above the other, then this conceptual representation is not built.

An example of a sentence where the reference object is conceptualized as an empty volume because it is not a well-defined high container is

(149) The boy is *in* the tree (the branches of the tree).
Le garçon est *dans* l'arbre.

See section 4.2.6.2 for a description of the effect of similar provisions and conditions.

This conceptualization is the one assumed in some of the sentences of the type described by Herskovits as "Physical object in the outline of another, or of a group of objects". Herskovits, however, does not differentiate between low and high. (See the last remark of section 4.2.6.1 for relevant discussion.)

4.2.6.5 Object is an institution one can be affiliated with and whose location is irrelevant

This section is meant to describe the English conceptual representation corresponding to the French one of section 4.2.1.12. The Provisional and Execution information are almost the same in this case as in the case of section 4.2.1.12, with the exception that here an extra Provision is that no article precedes the reference object. In this section, we simply recall the example of section 4.2.1.12.

(150) Joshua is two, he is not *in* school yet!
Joshua a deux ans, il n'est pas encore *à* l'école!

even though Joshua might happen to be located in the school when the sentence is uttered. We also recall that the English conceptualization of this case is empty volume, whereas the French one is point.

4.2.6.6 Object is a part of space or an environment

This section describes the English conceptual representation corresponding to the French one of section 4.2.1.13. Since the Provisional and execution information are exactly the same in this case as in the case of section 4.2.1.13, they are not repeated here. In this section, we simply recall the example of section 4.2.1.13.

(151) The boy is *in* the shade.
Le garçon est *à* l'ombre.

and recall that the English conceptualization of this case is surface, whereas the French one is point.

4.2.6.7 Object is a closed environment

In the provisional part, the following properties must be verified:

1. the object must have the natural function of creating an environment,
2. the object must have the natural function of hiding.

In the execution part, the conditions of the conceptual representation are the following:

1. the language considered can be English or French,
2. there is no condition involving the located object,
3. there are two highlighted properties of the reference object: it must be a part of space or environment, and can hide,
4. there is no extra constraint.

Condition 3 says that the relevant property of the reference object in this case is that it is a part of space or an environment.

An example of a sentence where the reference object is conceptualized as an empty volume because it is a closed environment is

(152) The boy is *in* the fog.
Le garçon est *dans* le brouillard.

In the provisional case, the unique Provision eliminates the case of section 4.2.1.13 in French, which treats the case of sentences of the type

(153) Le garçon est *à* l'ombre.
The boy is **at/in* the shade.

English does not distinguish between the two types of sentences (even though our system does).

This conceptualization is the one assumed in sentences of the type described by Herskovits as "Spatial entity in part of space or environment". However, Herskovits does not distinguish between the case where one can hide in the environment, and the case where one cannot hide in it. This is because English does not distinguish between the two cases.

4.2.6.8 Object is a piece of clothing

In the provisional part, the following provision must be verified:

The object must belong to the category 'clothing'.

In the execution part, the conditions of the conceptual representation are the following:

1. the language considered must be English, but not French,
2. the located object must be human,
3. the highlighted property of the reference object must be that it is a piece of clothing.
4. there is no extra constraint.

Condition 3 is self-explanatory, in this case.

An example of a sentence where the reference object is conceptualized as an empty volume because it is a piece of clothing is

(154) The man *in* the shirt ...
L'homme **dans/en* chemise ...

Note that in this case, French uses a different conceptualization: functional container (refer to case 4.2.10.4 for the corresponding French conceptual representation).

Note further that in French, in the case where the piece of clothing presents a particularity, the preposition used instead of *en* is *à/au*. For example, one says:

(155) L'homme *à* la chemise déchirée ...
The man *in* the ripped shirt ...

Our program, however does not take this last case into consideration (because it is very rarely used).

In the provisional part, the unique Provision simply characterizes the case.

In the execution part, Condition 2 rules out invalid sentences such as

(156) * The box is *in* the shirt.

because a box is not human. This conceptualization is the one assumed in sentences of the type described by Herskovits as: "Person in clothing".

4.2.6.9 Object is a throughway that can be obstructed

In the provisional part, the following provision must be verified:

The object must belong to the category throughway,

In the execution part, the conditions of the conceptual representation are the following:

1. the language considered must be English, but not French,
2. there is no condition involving the located object,
3. the highlighted property of the reference object must be that it is a throughway that can be obstructed,
4. there is no extra constraint.

Condition 3 says that the relevant property of the object in this case is that it is usually a throughway and we concentrate on the fact that it can be obstructed.

An example of a sentence where the reference object is conceptualized as an empty volume because it is a throughway is

(157) The tree *in* the road ...
L'arbre **dans le/en travers du* chemin ...

Note that in this case, French uses a different conceptualization: obstruction (refer to case 4.2.11.1 for the corresponding French conceptual representation).

In the provisional part, the unique Provision simply characterizes the case.

This conceptualization is the one assumed in sentences of the type described by Herskovits as: "Physical object obstructing roadway".

4.2.6.10 Object is air seen as a location

In the provisional part, the following provision must be verified:

The object must be the 'air',

In the execution part, the conditions of the conceptual representation are the following:

1. the language considered must be English, but not French,
2. the located object must be light or able to fly,
3. the highlighted property of the reference object must be that it is the air seen as a location,
4. there is no extra constraint.

Condition 3 is self-explanatory, in this case.

Examples of sentences where the reference object is conceptualized as an empty volume because it is air seen as a location are

(158) The bird *in* the air ...
L'oiseau **dans/en* l'air ...

(159) The plane is *in* the air ...
L'avion est **dans/en* l'air ...

Note that in this case, French uses a different conceptualization: functional container (refer to case 4.2.10.2 for the corresponding French conceptual representation).

In the provisional part, the unique Provision simply characterizes the case.

In the execution part, Condition 2 rules out sentences such as

(160) * The elephant is *in* the air.

This conceptualization is the one assumed in sentences of the type described by Herskovits as "Physical object in the air".

4.2.6.11 Object is a horizontal surface and has sides

This section describes the French conceptual representation corresponding to the English one of section 4.2.3.7. Since the Provisional and execution information are exactly the same in this case as in the case of section 4.2.3.7, they are not repeated here. In this section, we simply recall one of the examples of section 4.2.3.7.

(161) Maman est *dans* le bus.
Mom is **in/on* the bus.

and recall that the French conceptualization of this case is empty volume, whereas the English one is surface. (Note that this French case should be grouped together with the case of section 4.2.6.1, but is not in our implementation.)

4.2.7 Conceptualization as empty volume with a tiny width

In the conceptual representation described in this section, the node pointed to by the arc META contains the information specifying that the dimension of the metaphor is VOLUME, its fullness is EMPTY, and its width is TINY-WIDTH. This is a geometric conceptualization which indicates that the object considered is conceptualized as a container which can contain foreign objects, and that it has an almost non-existent width. The only type of objects conceptualized in this way are representations, such as pictures, paintings, etc.

4.2.7.1 Object is an environment and has a tiny width

This section describes the English conceptual representation corresponding to the French one of section 4.2.3.10. Since the Provisional and execution information are exactly the same in this case as in the case of section 4.2.3.10, they are not repeated here. In this section, we simply recall the example of section 4.2.3.10.

(162) The boy is *in* the picture.
Le garçon est **dans/sur* la photo.

and recall that the English conceptualization of this case is empty volume with a tiny width, whereas the French one is surface.

4.2.8 Conceptualization as a water body edge

In the conceptual representation described in this section, the node pointed to by the arc META contains the information specifying that the dimension[40] of the metaphor is EDGE-OF-WATER-BODY, its fullness and its width are NIL. This conceptualization is not purely geometrical: it involves other properties as well. It indicates that the object considered is conceptualized as the edge of a body of water. In this case, the conceptualization is a clear case of metonymy, where the entire object is conceptualized as only a part of itself. For example, in certain cases (4.2.8.1), a lake (the entire lake) can be conceptualized as only its edge (actually, only a part of its edge).

4.2.8.1 Object is a body of water whose edge is focused on

This section describes the French conceptual representation corresponding to the English one of section 4.2.2.3. Since the Provisional and execution information are exactly the same in this case as in the case of section 4.2.2.3, they are not repeated here. In this section, we simply recall the example of section 4.2.2.3.

(163) La maison **sur le /au bord du* lac ...
The house *on* the lake ...

and recall that the French conceptualization of this case is water body edge, whereas the English one is line.

4.2.9 Conceptualization as a solid-line

In the conceptual representation described in this section, the node pointed to by the arc META contains the information specifying that the dimension of the metaphor is EDGE-OF-OPEN-SPACE, its fullness and its width are NIL. This conceptualization is not purely geometrical. It indicates that the object considered is conceptualized as the edge of an open space. As above, in this case, the conceptualization is a clear case of metonymy.

[40] As mentioned earlier, the name 'dimension' is misleading: it presupposes that the information the arc named 'dimension' points to is a number. However, as we saw earlier, it can be a geometric object such as a surface, a volume, rather than a number. At this point the term is even more misleading, because the arc bearing it as a name points to information of a functional type as well as geometric.

For example, a park (the entire park) is, in some case (4.2.2.4), conceptualized as only its edge (actually, only a part of its edge).

4.2.9.1 Object is a solid geographic entity whose edge is focused on

This section describes the French conceptual representation corresponding to the English one of section 4.2.2.4. Since the Provisional and execution information are exactly the same in this case as in the case of section 4.2.2.4, they are not repeated here. In this section, we simply recall the example of section 4.2.2.4.

(164) La maison **sur le/en bordure du* parc ...
The house *on* the park ...

and recall that the French conceptualization of this case is solid-line, whereas the English one is line.

4.2.10 The conceptualizations as a functional container

In each of the conceptual representations described in this section, the node pointed to by the arc META contains the information specifying that the dimension of the metaphor is FUNCTIONAL CONTAINER, and its fullness and width are NIL. FUNCTIONAL CONTAINER is a conceptualization mid-way between a geometrical and a functional conceptualization. The geometrical conceptualization it is the closer to is EMPTY VOLUME, and the functional one it is the closer to is the functional meaning of a POINT [41]. The idea behind the term FUNCTIONAL CONTAINER is that the located object is located inside the reference object, but that there is an extra, closer link than containment between the objects involved in the relation. For example, in the case where the reference object is a piece of clothing, it does contain the located object (the person) but there is a closer relation between a person and his/her clothing than one of containment.

4.2.10.1 Object is a generic sea

This section describes the French conceptual representation corresponding to the English one of section 4.2.1.7. Since the Provisional and execution information are exactly the same in this case as in the case of section 4.2.1.7, they are not repeated here. In this section, we simply recall the example of section 4.2.1.7.

(165) Les conteneurs de marijuana sont déjà *en* mer.
The marijuana containers are already *at* sea.

and recall that the French conceptualization of this case is functional container, whereas the English one is point.

[41] Note that the conceptualization POINT should have been divided into two separate conceptualizations: one purely geometrical, and the second one functional.

4.2.10.2 Object is air seen as a location

This section describes the French conceptual representation corresponding to the English one of section 4.2.6.11. Since the Provisional and execution information are exactly the same in this case as in the case of section 4.2.6.11, they are not repeated here. In this section, we simply recall one of the example of section 4.2.6.11.

(166) L'oiseau *en* l'air ...
The bird *in* the air ...

and recall that the French conceptualization of this case is functional container, whereas the English one is empty volume.

4.2.10.3 Object is a bounded surface and a country or a continent

This section describes the French conceptual representation corresponding to the English one of section 4.2.4.1. Since the Provisional and execution information are exactly the same in this case as in the case of section 4.2.4.1, they are not repeated here. In this section, we simply recall the example of section 4.2.4.1.

(167) Le garçon est *en* Afrique.
The boy is *in* Africa.

and recall that the French conceptualization of this case is functional container, whereas the English one is empty volume.

4.2.10.4 Object is clothing

This section describes the French conceptual representation corresponding to the English one of section 4.2.6.9. Since the Provisional and execution information are exactly the same in this case as in the case of section 4.2.6.9, they are not repeated here. In this section, we simply recall the example of section 4.2.6.9.

(168) L'homme *en* chemise ...
The man *in* the shirt ...

and recall that the French conceptualization of this case is functional container, whereas the English one is empty volume.

4.2.11 Conceptualization as an obstruction

In the conceptual representation described in this section, the node pointed to by the arc META contains the information specifying that the dimension of the metaphor is OBSTRUCTION, and its fullness and width are NIL. In this case, the conceptualization is mainly functional. The reference object is seen as an object that is obstructed, that is, an object which normally allows passage, but that, in the circumstances of the case, is blocked. The located object in this conceptualization is assumed to be the obstructor.

4.2.11.1 Object is a throughway [This case is not implemented in the system]

This section is meant to describe the French conceptual representation corresponding to the English one of section 4.2.6.10. Since the Provisional and execution information are exactly the same in this case as in the case of section 4.2.6.10, they are not repeated here. In this section, we simply recall the example of section 4.2.6.10.

(169) L'arbre *en travers du* chemin ...
The tree *in* the road ...

and recall that the French conceptualization of this case is obstruction, whereas the English one is empty volume.

4.2.12 Conceptualization as an intersection

In the conceptual representation described in this section, the node pointed to by the arc META contains the information specifying that the dimension of the metaphor is INTERSECTION, and its fullness and width are NIL. In this case, the conceptualization is purely geometric. It is the point of intersection of two lines.

4.2.12.1 Object is a point located on a linear object

This section describes the French conceptual representation corresponding to the English one of section 4.2.1.2. Since the Provisional and execution information are exactly the same in this case as in the case of section 4.2.1.2, they are not repeated here. In this section, we simply recall the example of section 4.2.1.2.

(170) La station service est *à la jonction* avec l'autoroute.
The gas station is *at* the freeway.

and recall that the French conceptualization of this case is intersection, whereas the English one is point.

4.2.13 Conceptualization as an opposition

In the conceptual representation described in this section, the node pointed to by the arc META contains the information specifying that the dimension of the metaphor is OPPOSITION, and its fullness and width are NIL. In this case, the conceptualization is physical: an object is opposing a force on the object being conceptualized.

4.2.13.1 Object is a wall which can be contiguous to other objects

This section describes the French conceptual representation corresponding to the English one of section 4.2.3.8. Since the Provisional and execution information are exactly the same in this case as in the case of section 4.2.3.8, they are not repeated here. In this section, we simply recall the example of section 4.2.3.8.

(171) *Contre* le mur de gauche, se trouve une commode.
On the left wall, there is a chest of drawers [Herskovits (1986) p.145].

and recall that the French conceptualization of this case is opposition, whereas the English one is surface.

4.3 Some of the points not considered by the system

Not all cases were taken into consideration in this study. We concentrated on the most common ones, leaving out the peripheral ones. Our main source of inspiration is Herskovits's case studies [Herskovits (1986), chapter 9]. We did not, however treat all the cases which she studied. In particular, three of her cases were not taken into consideration: "Physical object at a distance from point, line, or plane", "Accident/object part of physical object", and "Accident/object part of physical or geometric object". In the first case, our grammar did not allow such constructions, and in the second and third cases, we did not find the case that clear in French.

As well, within Herskovits's cases, some exceptional cases were not taken into consideration. Such cases are, for example,

(172) The boy is *on* the third floor.
Le garçon est **sur le/au* troisième etage.

(173) The fruit is *on* the bottom of the yogurt
Les fruits sont **sur le/au* fond du yogurt

(174) The dog *on* a leash ...
Le chien **sur la/en* laisse ...

(175) The man is *in* prison ...
L'homme est **dans la/en* prison ...

Actually, these cases might not be exceptional, but their relations to other cases are not straightforward. As well, they seemed isolated, and we preferred to study other more important cases, instead.

Not all cases presented in this chapter were implemented (all such are indicated), due to their special grammatical needs or their peripheral nature. In addition, a point not considered by our implementation is that sometimes including the verb 'to be' in a sentence makes it sound better, and sometimes the opposite case occurs (in fact, in certain cases, the inclusion or exclusion of 'to be' is simply ungrammatical). Our system does not differentiate the bad from the good sentences in these terms, and simply accepts them all. This is due to the fact that we did not try to explain what makes some of these sentences sound good, and what does not. We, however, used examples, which we thought sounded better, such as,

(176) The house *on* the lake ...

rather than

(177) The house is *on* the lake ...

and

(178) The boy is *at* the supermarket

rather than

(179) The boy *at* the supermarket ...

Note that, such a problem seems interesting because in the last set of examples, sentence (178) seems to imply more naturally that the boy went shopping, whereas sentence (179) seems to imply naturally that the boy works there (unless a context is given that clarifies the case).

The last point that we will discuss here relates to footnote no. 20 of this chapter. We mentioned that we intended to group together similar cases within each language, but did not carry out this operation systematically. We only attempted it in a few cases (geographic entities and two of the cases of environments or part of spaces).

4.4 Comparisons between French and English

In this section, we will first summarize the differences between English and French that were found in section 4.2. We will then try to give a succinct analysis of these differences.

4.4.1 Summary of the differences between English and French

We distinguished 15 cases in which the translation of English sentences into French diverges from the most common pattern in which

'at' is translated as 'à/au'

'on' is translated as 'sur'

'in' is translated as 'dans'

In six of these cases, the preposition is translated into another of the three basic topological prepositions. In the nine other cases, the preposition is translated into another preposition than 'in', 'on', or 'at'. The first six cases follow:

'on' is translated by 'à/au' in the case of sections 4.2.1.9 and 4.2.3.6, where the object is involved in a passive support relation. An example of this case is

(180) The lamp is *on* the ceiling.
Le lustre est *au* plafond.

'on' is translated by 'dans' in the case of sections 4.2.3.7 and 4.2.6.12, where the object is an horizontal support and has sides. An example of this case is

(181) The boy is *on* the bus.
Le garçon est *dans* le bus.

'in' is translated by 'au' in the case of sections 4.2.1.10 and 4.2.4.1, where the object is a country or a continent (and is masculine in French) An example of this case is

(182) The boy is *in* Canada.
Le garçon est *au* Canada.

'in' is translated by 'à/au' in the case of sections 4.2.1.12 and 4.2.6.5, where the object is an institution one can be affiliated with and need not be actually located in. An example of this case is

(183) Joshua is two, he is not *in* school yet!
Joshua a deux ans, il n'est pas encore *à* l'école!

(Even though Joshua could be actually located in the school when the sentence is uttered.)

'in' is translated by 'à/au' in the case of sections 4.2.1.13 and 4.2.6.6, where the object is a part of space or an environment. An example of this case is

(184) The boy is *in* the shade.
Le garçon est *à* l'ombre.

'in' is translated by 'sur' in the case of sections 4.2.3.10 and 4.2.7.1, where the object is an environment and has a tiny width. An example of this case is

(185) The boy is *in* the picture.
Le garçon est *sur* la photo.

The preposition is translated into a non-basic topological preposition in the nine following cases:

'in' is translated by 'en travers de' in the case of sections 4.2.6.10 and 4.2.11.1, where an object is a throughway. An example of this case is

(186) The tree *in* the roadway ...
L'arbre *en travers du* chemin ...

'at' is translated by 'à l'intersection avec' in the case of sections 4.2.1.1 and 4.2.12.1, where the object is a point located on a linear object. An example of this case is

(187) The gas station is *at* the freeway.
La station service est *à* l'intersection avec l'autoroute.

'at' is translated by 'en' in the case of sections 4.2.1.7 and 4.2.10.1, where the object is a generic sea. An example of this case is

(188) The marijuana containers are already *at* sea.
Les conteneurs de marijuana sont déjà *en* mer.

'on' is translated by 'au bord de' in the case of sections 4.2.2.3 and 4.2.8.1, where the object is the edge of a body of water. An example of this case is

(189) The house *on* the lake ...
La maison *au bord du* lac ...

'on' is translated by 'en bordure de' in the case of sections 4.2.2.4 and 4.2.9.1, where the object is the edge of a solid geographic entity. An example of this case is

(190) The house *on* the park ...
La maison *en bordure du* parc ...

'on' is translated by 'contre' in the case of sections 4.2.3.8 and 4.2.13.1, where the object is a wall. An example of this case is

(191) The chest of drawers is *on* the left wall.
La commode est *contre* le mur de gauche.

'in' is translated by 'en' in the case of sections 4.2.4.1 and 4.2.10.3, where the object is a country or a continent (and is feminine or is masculine and starts with a vowel). Two examples of this case are

(192) The boy is *in* France.
Le garçon est *en* France. [France is feminine]

(193) The boy is *in* Afghanistan.
Le garçon est *en* Afghanistan. [Afghanistan is masculine]

'in' is translated by 'en' in the case of sections 4.2.6.9 and 4.2.10.4, where the object is clothing. An example of this case is

(194) The man *in* the red shirt ...
L'homme *en* chemise rouge ...

'in' is translated by 'en' in the case of sections 4.2.6.11 and 4.2.10.2, where the object is air as a location. An example of this case is

(195) The bird *in* the air ...
L'oiseau *en* l'air ...

4.4.2 Analysis of the differences

In this section, we will try to generalize and analyze the differences between English and French that were summarized in section 4.4.1. We first do a quantitative analysis of the results, and then a qualitative one. Our conclusions are not very rigorous, given the fact that we did not gather many examples; therefore, they are simply tentative ones. The reason that we included them is to indicate trends that we noticed.

Quantitative analysis of the differences

We listed the differences in terms of their uses with respect to their corresponding preposition in the other language. We conclude that:

- In 8 of the cases where 'in' is used in English, 'dans' is not used in French
- In 4 of the cases where 'on' is used in English, 'sur' is not used in French
- In 2 of the cases where 'at' is used in English, 'à/au' is not used in French
- In 4 of the cases where 'à/au' is used in French, 'at' is not used in English
- In 1 of the cases where 'dans' is used in French, 'in' is not used in English
- In 1 of the cases where 'sur' is used in French, 'on' is not used in English
- 'en' is used in 4 French cases
- *miscellaneous prepositions* are used in 4 French cases

Tentative conclusions of the quantitative analysis

- 'à/au' is used in many more cases that 'at'.
- 'in' is used in many more cases than 'dans'.
- 'on' is used in a few more cases than 'sur'.
- 'en' is an important preposition in the French expression of topological situations. It might be counted as part of the French set of basic topological prepositions.

Qualitative analysis of the differences

The differences between English and French might be grouped in three different trends:

First, some uses of prepositions are simply conventional. This is the case of the use of 'au' or 'en' with the name of geographical entities in French expressions: the use is simply due to syntax.

Second, the English/French pairs of corresponding prepositions might have different meanings in some cases. This seems true in the following cases: "Object is an institution understood as a place one is affiliated with and whose location is irrelevant" (sections 4.2.1.12 and 4.2.6.5), and "Object is a throughway" (sections 4.2.6.9 and 4.2.11.1). In the first two cases, 'in' takes a functional meaning that 'dans' does not have: the idea of affiliation. In the third case 'in' takes another functional meaning that 'dans' does not have: the idea of obstruction.

Finally, in some bordering cases, where several conceptualizations are cognitively acceptable, languages might "choose" their interpretation. In our study, we considered

this "choice" arbitrary, but we believe that some general conceptual rules about languages must exist that direct this "choice" (see chapter 7). This applies in all the cases that use different conceptualizations in English and French, and which are not explained by any of the two previous analysis. We will illustrate this analysis with one example:

(196) The man is *in* the picture.
L'homme est *sur* la photo.

In English, the picture is conceptualized as a container, i.e., even though the width is thin, it is considered relevant. The man therefore appears to be 'in' the picture. In French, the width is seen as so thin that it is considered nonexistent. The man, therefore appears to be 'on' the picture. Both interpretations are cognitively acceptable but a different one is preferred in each language.

4.4.3 Tentative conclusions about the qualitative analysis

The first point agrees with Herskovits's view that some of the uses of the topological prepositions are simply conventional.

From the second point, we conclude that 'in' seems to have a functional meaning that 'dans' does not have, i.e., 'dans' is purely locative. We attempt to explain the use of the extra preposition 'en' in French by this fact. We observed that 'en' stands between 'dans and 'à' in meaning, since 'en' is more functional than 'dans' but less than 'à', and more locative than 'à' but less than 'dans'. It might be because of the lack of functional meaning of 'dans' that an extra preposition, more functional than 'dans', but less so than 'à' was needed in French, the preposition 'en'.

From the third point, together with the conclusions of the quantitative analysis, we can try to derive general points about the two languages: these might be that since when a choice is possible, 'à/au' is chosen more often than any other prepositions, and since 'à/au' is less precise than 'dans', 'sur', or their equivalents in English, then French might be considered as less precise[42] as topological relations involving basic topological prepositions.

[42]This conclusion, however, is very quick, given that 'à/au' is used in only a few more cases in French than 'at', in English. We, decided to make it, however, because we observed that the few cases in which 'à/au' is used in French where 'at' is not are important cases of locative expressions.

Chapter 5

Error and ambiguity handling

5.1 Introduction

This chapter describes two extra tasks performed by the system that were not described in chapter 3: error and ambiguity handling. The system responds to the cases of error or abnormality by issuing a message and aborting the run. Cases of ambiguity are handled by considering all acceptable meanings of the sentence and processing each of them separately. In section 5.2 below, errors are subdivided into several types and the method of detection and handling is described for each type. Section 5.3 discusses ambiguity in a similar manner.

5.2 Error detection

The system detects two types of error: errors at the lexical level, and errors at the conceptual level.

5.2.1 Lexical errors

We call the misspelling of a word, or the fact that it is absent from the lexicon a *lexical error*. When such an error is detected, the program issues an error message and aborts. This type of error is detected by the SLU. In the following example,

(197) * The booy is in the shade.

'booy' is not defined in the lexicon.

5.2.2 Conceptual errors

We use the label *conceptual error* for three types of errors:

- Errors in the conceptualization of the reference object
- Errors in the use of a conceptualization of the reference object
- Abnormality of certain uses

The first type of error is detected by the DEC, before the located object is even considered and the second and third types of error or abnormality are detected by the SB, once the located object has been considered. Note that even though we treat the third case together with the other two, we acknowledge its difference: it is not an error per se, but rather an abnormality that could be acceptable in special contexts.

Errors of conceptualization

Errors of conceptualization occur when the preposition requires the reference object to be conceptualized in a way that it cannot be in the language considered. An example of a sentence where such an error occurs is:

(198) * The boy is at the shade.[1]

This sentence is erroneous because 'at' requires 'shade' to be conceptualized as a point, but 'shade' used as a reference object can never be conceptualized as a point in English. Such an error is detected at the DEC level, when possible conceptual representations of the reference object must be chosen in accordance with the requirements of the preposition used in the sentence. The DEC decides that there is such an error if the set of conceptual representations it finds is empty. An error message is then issued, saying that no conceptualization was found for the reference object that satisfies the requirements of the preposition, and the run is aborted.

Usage errors

Usage errors occur in expressions in which the demands of the preposition are satisfied by the reference object, but the conditions required of the located object by the conceptual representation considered are not. These conditions may or may not also involve the reference object. They are stored in the node pointed to by the LOCOBJCONDITION arc of the conceptual representations of the reference object.

Usage errors are detected by the SB, prior to building an objective scene of the sentence. Such an error prevents the system from building the objective scene it was considering when the error was detected, but does not result in the abortion of the run, for even if such an error is detected during an attempt to build a scene, another scene may be perfectly acceptable. If, however, no scene can be built, either because errors of the type described here and/or normality errors of the type described in the next section are detected, then a message stating that no scene could be built is issued and the run is aborted.

We illustrate with examples the case where the run is not aborted and the case where it is. Consider the following:

(199) The box is on the bed.

[1] Note that in French, the word for word translation of this sentence, taking 'at' to be 'à', is perfectly correct.

For this sentence, after the DEC has operated, two scenes are attempted. The first one is a scene where the highlighted property of the reference object is that it is part of an object it can support, and the second one is a scene where the highlighted property of the reference object is that it is a planar horizontal support. The first scene cannot be built because a usage error is detected. This error is detected because the condition concerning the located object that must be satisfied for the scene to be built, that the reference object be part of the located object, is not satisfied: a bed is not part of a box (note that this scene would be built if the reference object was, for example, 'its top', because 'The box is on its top' is acceptable). However, the second scene that is attempted is perfectly valid, and the run is not aborted. Now consider the following sentence:

(200) The man is in the table.

For this sentence, after the DEC has operated, only one scene can be attempted. In this scene, the highlighted property of the reference object is that the object is apprehended in its whole. The condition involving the located object in which such a scene can be built is that the located object is allowed in the reference object. Since a man is not a constituent of a table and cannot be embedded in a table, this property is not verified. Given that there is only one scene to be attempted, the type of error described in this section causes the run to abort.

The abnormalities

Abnormalities occur in expressions in which the demands of the preposition are satisfied by the reference object, the conditions required of the located object by the conceptual representation of the reference object are satisfied as well, but the normality conditions are not. We use the same idea of normality as Herskovits's [Herskovits (1986), p. 20]:

> First, a normal situation conforms to the laws of physics — the common sense physics of ordinary solid objects, liquids and gaseous substances ...Second, objects are where they belong — most of them near the earth, within the field of gravity ...Finally, the objects are "normal", interact according to their normal function, and people interact with them in normal fashion.

The difference between the case of abnormality and the case of usage errors is that in the former, the non-satisfaction of the conditions involving the located object causes an error, whatever the context may be, whereas, in the latter, even though in the normal context the sentence sounds wrong or at least strange, there might be contexts in which it is acceptable (see the examples at the end of the section.) The way in which the abnormality detection procedure works is similar to the way in which the usage error detection procedure does, the only difference being that the condition checked in the usage error case is stored with the conceptual representation of the specific reference object considered, whereas the condition checked in the abnormality case is stored procedurally in the system, and is common to all the scenes, but for a few exceptions also listed procedurally in the system. Our reason for choosing this type of implementation was as

follows: since the normality conditions apply to most of the scenes, it was easier to list the exceptions than to list all the cases in which they apply. We stored two conditions to be checked centrally: the checking of the respective sizes of objects and the checking of the materiality of the located object. The size rule states that if the scene attempted to be built is not one of the exceptions, then the located object must be smaller than (or if the located object is small, the same size as) the reference object. The materiality rule says that if the scene attempted to be built is not one of the exceptions, then the located object must be material. By material, we mean that the object must be palpable. Gas or atmospheric entities are not material; solids or liquids are. The exceptions for size checking are the following scenes: "physical object on part of itself" and "gap/object embedded in physical object". The exceptions for materiality checking are the following scenes: "physical object over a geographic object", "spatial entity contained in lower part of physical object", "spatial entity contained in higher part of physical object", "spatial entity in area", and "gap/object embedded in physical object". Here are two examples of sentences in which the size and materiality checking should be satisfied, but one of them is not:

(201) * The elephant is on the flower.

and

(202) * The shadow is on the bus.

In the first case, the size (and weight) rule is not respected, and in the second case, the materiality rule is not respected. We now illustrate our distinction between abnormality and errors, using examples (200), (201), and (202). In a world where flowers are giants and elephants can fly, there is no problem with sentence (201), and in a world where shadows are personified, sentence (202) is acceptable too. However, no world can be created in which sentence (200) is acceptable, and sentence (200) is erroneous, rather than abnormal.[2]

5.3 Ambiguity detection

When the system detects ambiguity it considers one by one each possible context in which the sentence might have been uttered, and provides a solution for each of them. If our system dealt with paragraphs, a technique for resolving ambiguity could have been implemented. However, in the case of isolated sentences, a system has only three possibilities regarding ambiguity: it can simply ignore it, interact with the user by prompting him/her to specify the context in which the sentence was uttered, or not interact with the user but consider each possible context one after the other, and provide a solution for

[2]This distinction, however, is problematic: our claim about sentence (200) is not totally justified: there might be a world in which sentence (200) is acceptable. In fact, in a context were a man was captured by a giant who used him as a nail and stuck him in its giant table, sentence (200) is acceptable. By considering other such sentences, which we first thought were erroneous, we were able to make up contexts in which these sentences were acceptable. These contexts are, however, "more abnormal" than the abnormal ones, in that they were not obvious right away. This notion, however, is subjective.

each of them. The last two ways are equivalent in level of sophistication. We opted for one of these: the non-interactive method. This method was chosen over the interactive one because interaction with the user would have slowed down the (real) running time of the system.

The system detects two main types of ambiguity: lexical ambiguity and conceptual ambiguity. For both types, the ambiguity can be simple or complex. In order for all the possible ambiguities to be satisfactorily dealt with, according to our method, a special way of identifying the entities built by the program was needed: the three parameters introduced in section 3.5, SENTENCEID, LEXICALMEANINGSID, and CONCEPTUALMEANINGID. SENTENCEID indicates the identity of the sentence being processed, LEXICALMEANINGSID indicates the French combination of the located and reference objects being used (this was designed to cope with lexical ambiguity), and CONCEPTUALMEANINGID distinguishes the different scenes built for the same sentence, from one another (this was designed to deal with conceptual ambiguity).

5.3.1 Lexical ambiguity

We use the term *lexical ambiguity* for the case where the located and/or the reference object(s) have one English name and several French ones. The system detects and handles two cases of lexical ambiguity: the simple and the complex cases.

Simple lexical ambiguity

Simple lexical ambiguity is ambiguity where one English word has several meanings in French but these different meanings do not affect the spatial relation. We assumed that the number of French words having one same equivalent in English would rarely, if ever, be greater than two and we designed our system accordingly. One example of simple lexical ambiguity follows:

(203) The *cat* is in the room.

where this sentence can be translated as

(204) *Le chat* est dans la chambre.

or

(205) *La chatte* est dans la chambre.

The only difference between the two versions is the gender of the cat ('chat' is masculine and 'chatte' is feminine). This difference does not influence the choice of preposition. Note that the ambiguous word could also be the reference object, and that the relation would not be affected either, as in:

(206) She stuck a sticker on *the cat.*

which is translated as

(207) Elle a collé un auto-collant sur *le chat.*

or

(208) Elle a collé un auto-collant sur *la chatte.*

Our program outputs both translations — given that no extra-sentential context is available that could help disambiguate the sentence. (The only extra-sentential context is the normality of the scene, assumed true in our system, and which was discussed in section 5.2. Normality, however, does not suffice here for disambiguating the sentence.) The way the simple type of lexical ambiguity is handled is by simply specifying the two possible French translations of the English ambiguous nouns in the lexicon. In the knowledge base, the concept of a cat is represented by a unique node. Thus, the system has different words of the same language in its lexicon for a single concept. We could use this technique because the gender of a cat does not affect the preposition used in the spatial relation.

Complex lexical ambiguity

Complex lexical ambiguity is lexical ambiguity of the located or reference object that affects the choice of preposition. The rule concerning the number of possible such translations of a noun that was adopted here, is the same as the one of the previous case. An example of the type of complex lexical ambiguity follows:

(209) The boy is *on the street.*

which might be translated as

(210) Le garçon est *dans la rue.*

or

(211) Le garçon est *sur la chaussée.*

The first translation means that the boy is either on the sidewalk or on the street itself; the second one specifies the roadway meaning of a street. English has only one noun where French has two: 'rue', which is thought of as the roadway together with the houses surrounding it and 'chaussée', which is thought as the roadway alone. In the case of 'rue', the preposition 'dans' must be used in topological descriptions and in the case of 'chaussée', it is 'sur' that must be used. This type of ambiguity is handled by our system by having two different concepts of street in the knowledge base, one of which is expressed as 'rue' in French, and the other of which is expressed as 'chaussée' in French. The lexicon indicates at the 'street' entry that it is a conceptually ambiguous noun, with two conceptually different translations in French. Lexically ambiguous sentences of the complex type are processed as if two different sentences were input, and the two possible French translations are output (the features needed for this purpose are described in the next section).

5.3.2 Conceptual ambiguity

We use the term *conceptual ambiguity* in the case where a single sentence can be conceptualized in several ways. The system detects and handles two types of conceptual ambiguity: simple and complex.

Simple conceptual ambiguity

In the simple type of conceptual ambiguity, a single English sentence can be conceptualized in several different ways. The sentence is translated into a single French sentence that is ambiguous in the same way. This type of ambiguity differs from the preceding type in that each word is translated in only one way in French. An illustration of this type of ambiguity follows:

(212) The boy is *at* the supermarket.

can be understood as "the boy is shopping at the supermarket", or "the boy is on a trajectory going by the supermarket, and is currently located at the supermarket". Its French translation is

(213) Le garçon est *au* supermarché.

which carries the same ambiguity as the English sentence. No special features were provided to the lexicon for dealing with this case. Two features were, however, provided to the SB and the main module: the SB was provided with a loop allowing multiple scenes to be built for the same sentence. The main program was provided with a loop calling each of the modules following the SB as many times as scenes were built.[3] (Note however that these loops were also needed in the previous case, where multiple scenes were also built.)

Complex conceptual ambiguity

Complex conceptual ambiguity is similar to simple conceptual ambiguity. The only difference is that, whereas in the case of simple conceptual ambiguity the French sentence carries the same ambiguity as the English sentence, in this case, the ambiguity is not carried into the translation; the sentence is translated into two different sentences, rather than one. An example of complex conceptual ambiguity follows:

(214) The boy is *on* the bus.

This sentence is ambiguous in that the boy can be located on the roof of the bus, or can be riding the bus. This sentence is translated into two French sentences, one for each case. The sentence is translated as

(215) Le garçon est *sur* le bus.

in the case where the boy is located on the roof of the bus, and as

[3]The loop of the main program is shown in the outline of the algorithm in chapter 3: the inner loop. The outline is not detailed enough to show the loop of the SB.

(216) Le garçon est *dans* le bus.

in the case where the boy is riding the bus. Note that in the cases of conceptual ambiguity, it is the preposition which is ambiguous in English, whereas in the case of lexical ambiguity, it is the located or the reference objects that are ambiguous.

No special feature was added for handling the case of complex conceptual ambiguity: it is automatically taken care of by the same features as the case of simple conceptual ambiguity.

In the cases where it is English that is lexically or conceptually ambiguous, then nothing needs to be done: in the cases of lexical ambiguity, two different English nouns have the same French translation in the lexicon; in the case of conceptual ambiguity, the English sentence is translated into an ambiguous French sentence, the preferred meaning presumably being clear in context.

Chapter 6

Sample translations

6.1 Introduction

The examples presented in this chapter demonstrate the performance of the system. To repeat, the main goal of the system is to translate English sentences or expressions involving one of the basic topological prepositions into French equivalents. The system is also able to identify and deal with ambiguities and errors in the source sentence, as was discussed in chapter 5. The runs presented below show translations of non-ambiguous, ambiguous, and erroneous input sentences. In section 6.2, we will explain the format of the output; section 6.3 shows runs of ambiguous or non-ambiguous valid sentences; section 6.4 demonstrates our handling of invalid input sentences; and section 6.5 summarizes these results.

6.2 Explanation of the format of the output

We will now explain what the information output by the system indicates. The format of the output matches the description of the system given in chapter 3. We will illustrate this section with the example of figure 6.1, the sentence "The boy is at the tree".

Input

The first line of the output shows the English sentence input to the system.

Initialization

The section called INITIALIZATION corresponds to the phase of the same name in chapter 3.

The first paragraph of this section is a report on the lexical ambiguity of the located and reference objects. This report was derived, together with other information not shown here, by the SLU. In figure 6.1, this paragraph is only one line long and states that the located and reference objects are not lexically ambiguous.

```
 : The boy is at the tree

***** INITIALIZATION *****

The located and the reference objects are not lexically ambiguous.

In English, 'tree', when used as a reference object, can be conceptualized as:
 C1. a POINT if its highlighted property is one of the following:
    1. landmark-in-highlighted-medium
 C2.  SURFACE if its highlighted property is one of the following:
    1. part-of-object-support
    2. non-horizontal-object-involved-in-active-support-relation
    3. non-horizontal-object-involved-in-passive-support-relation
 C3. a FULL VOLUME if its highlighted property is one of the following:
    1. object-apprehended-as-a-whole
 C4. an EMPTY VOLUME if its highlighted property is one of the following:
    1. well-defined-low-container
    2. non-well-defined-high-container


**** DERIVATION OF THE OBJECTIVE MEANING ****

Given the ideal meanings of 'at', only set C1 is selected.

For scene SC1.1 to be built, the following conditions must be satisfied.
 A. The located object must be mobile
       * This condition is verified because a boy is mobile.
 B. The situation must be normal
       * The size check is satisfied.
       * The materiality check is satisfied.
Because all the conditions are verified, scene SC1.1 is built.

***** TRANSLATION INTO FRENCH *****

The description of scene SC1.1 follows:
'Object at landmark-in-highlighted-medium'
Both English and French use conceptual representation C1.1 to express this scene.
In addition, the context should be the following:
the-located-object-must-be-on-a-defined-trajectory
The French translation of the sentence follows:
 : Le garcon  est  a l' arbre

 CPU time : 107.66     GC time : 0.00
```

Figure 6.1: The run of "The boy is at the tree"

The next part of this section shows all the English conceptual representations of the reference object built by the COB. We print only the information contained in the conceptual representation of objects which we judge relevant for the comprehension of the algorithm. C1, C2, etc. are the types of metaphors built, that is, the information pointed to by the META arc of the conceptual representation (POINT, SURFACE, LINE, EMPTY VOLUME, etc.). In figure 6.1, C1 is POINT, C2, SURFACE, C3, FULL VOLUME, and C4, EMPTY VOLUME. In the list below each Cx, the items show the properties of the reference object that must be highlighted for a conceptualization of type Cx to be instantiated. Each item represents the set of highlighted properties of one conceptual representation. To summarize, C$x.y$ is a conceptual representation whose META arc points to the metaphor specified in the output on the line starting with 'Cx.', and whose CONDITION arc points to a node whose emanating HIGHLIGHT arc points to the information specified on the first line following the line Cx, that starts with 'y.'. In figure 6.1, C1.1, for example, is a conceptual representation whose META arc points to POINT and whose CONDITION arc points to a node whose emanating HIGHLIGHT arc points to 'landmark in highlighted medium'.

Derivation of the objective meaning

The section called DERIVATION OF THE OBJECTIVE MEANING corresponds to the phase of the same name in chapter 3.

The first line is output by the DEC. It indicates, on the basis of the ideal meanings of the preposition used in the sentence, which metaphors can possibly be in use in the sentence. In figure 6.1, it is stated that given the ideal meaning of 'at', only C1 can be selected. The output does not explain how the DEC proceeds because it does it in a straightforward operation which simply consists in matching the parts of the ideal meanings of the preposition indicating the conceptualization it requires with the set of conceptualizations of the reference object built by the COB. In figure 6.1, 'at' has a unique ideal meaning: 'Object contiguous with a POINT'. The DEC matches this ideal meaning with C1 whose conceptualization is POINT and eliminates C2, C3, and C4, whose conceptualizations are SURFACE, FULL VOLUME, and EMPTY VOLUME, respectively.

The remainder of this section is output by the SB. The information of the SB is output in the form of paragraphs separated from one another by a blank line. In figure 6.1, only one paragraph is output. Each paragraph shows the main operations in the process of building one objective scene. As well, the paragraph contains an indication of the success or the failure of each operation, and of the entire process. More precisely, scene SC$x.y$ represents the type of scene (or attempted scene) whose reference object is conceptualized by the metaphor of the conceptual representation C$x.y$, in English. In figure 6.1, only scene SC1.1 is attempted. For a scene SC$x.y$ to be built, at most two types of conditions must be satisfied: the condition involving the located object and the normality condition. The condition involving the located object is the one pointed to by the arc LOCOBJCONDITION of the conceptual representation C$x.y$ (in some case, LOCOBJCONDITION points to NIL, which means that no condition involving the located object

requires to be checked). The normality conditions are the size and materiality conditions discussed in chapter 5 (here again, all the conditions do not need to be checked). Each of the two types of conditions is stated as condition A or B and the conclusion of the checking is printed, preceded with a '*' in the case where the condition is verified or does not matter, and with an 'o', in the case where the condition is not verified. In figure 6.1, the condition involving the located object is that it must be mobile. This condition is verified. The two normality conditions are verified as well. The last line or two lines of each paragraph indicate if the scene is built or not, and if not, that the conceptual representation corresponding to the scene is discarded. In figure 6.1, it is stated that scene SC1.1 was built.

Translation into French

The last section, TRANSLATION INTO FRENCH corresponds to the sections *Derivation of the target-language preposition* and *Finalization* of chapter 3. Again, this section is organized in paragraphs separated by a blank line. In figure 6.1, only one paragraph is needed.

The first three-to-six lines (depending on whether or not the conceptualization used in English and French is the same, and whether or not extra-sentential information is provided) of each paragraph are output by the MATCHER; the output of the FPD is not shown because it is simply the French preposition that must be used in the current case, which is shown later; and the last two lines are output by the Finalizer. The output of the MATCHER shown is first, the type of the scene chosen — this information is output with the goal of helping the reader understand the meaning of the sentence intended in the translation provided in the particular paragraph considered; second, a mention indicating whether English and French use the same metaphor or not in order to describe the same situation, and if not, what metaphor is used in French; and third, when needed and available, the extra-sentential context in which such a sentence is valid. In figure 6.1, the type of scene chosen is: 'Object at landmark in highlighted medium', the same conceptualization is used in English and in French: C1.1, and the extra-sentential context is: 'The located object must be on a defined trajectory'. The output of the Finalizer is simply the translation of the input sentence into French. In figure 6.1, the translation is: "Le garçon est à l'arbre".

Run-time

Finally, the last line is output by SNePS, and indicates the CPU time and the GC (Garbage Collection) time used by our system to execute the run printed. Note, however, that the CPU time does not indicate a reliable value. This is because during each run, many nodes are built that slow the running time of the system down. At points when this running time is getting much too slow, we reset the network. In light of this, taking the CPU time as a point of comparison between sentences, is not meaningful, since the running time depends principally on the conditions of the network, rather than on the sentence alone. We show the CPU time values only to give the reader an impression of the time it takes the system to translate a sentence.

Note, finally, that in all the examples shown, certain information has been deleted from the actual output of the system in order to save space. These are the messages issued by the module loading the GATN grammar and the commands calling and exiting from the GATN parser.

6.3 Example of valid sentences

In this section, we will show examples of runs. These runs are organized by types of metaphor used in English sentences. The sentences processed here might be lexically ambiguous, conceptually ambiguous, or non-ambiguous. Erroneous sentences will be presented in section 6.4. The information stating whether a sentence is ambiguous or not is indicated in the output, and the ambiguous sentences are treated as described in chapter 5. We present the runs case by case and in the order in which these cases were presented in chapter 4. Not all the cases of chapter 4, however, are presented. Each of the runs presented in this section handles its task satisfactorily and outputs the expected translation. Unsuccessfully handled cases are discussed in section 6.5.

6.3.1 The conceptualizations as a point

This section shows runs of sentences that belong to cases described in section 4.2.1 of chapter 4.

6.3.1.1 Reference object is a landmark in a highlighted medium

The examples of this section show runs of sentences belonging to section 4.2.1.3 of chapter 4. The conceptualization used in this case is the same in English and French.

Our first example, a simple instance of the case, was already shown in figure 6.1: "The boy is at the tree".

The second example, "The boy is at the factory", is more complicated and demonstrates how conceptual ambiguity is ruled out by the system: the sentence is not conceptually ambiguous because, in the knowledge base of the system, a boy *cannot* work at a factory[1] and the scene 'Object at location with a purpose' is discarded. In fact, in the case where 'factory' is replaced by 'supermarket', the sentence is ambiguous, because a boy *can* shop at a supermarket (second example of section 6.3.1.2).

```
 : The boy is at the factory

***** INITIALIZATION *****

The located and the reference objects are not lexically ambiguous.

In English, 'factory', when used as a reference object, can be conceptualized as:
 C1. a POINT if its highlighted property is one of the following:
    1. landmark-in-highlighted-medium
    2. location-with-a-purpose
 C2.  SURFACE if its highlighted property is one of the following:
    1. plane-horizontal-support
    2. part-of-object-support
 C3. an EMPTY VOLUME if its highlighted property is one of the following:
    1. well-defined-low-container


**** DERIVATION OF THE OBJECTIVE MEANING ****

Given the ideal meanings of 'at', only set C1 is selected.

For scene SC1.1 to be built, the following conditions must be satisfied.
 A. The located object must be mobile
       * This condition is verified because a boy is mobile.
 B. The situation must be normal
       * The size check is satisfied.
       * The materiality check is satisfied.
Because all the conditions are verified, scene SC1.1 is built.
```

[1] Note that this is a simplification of the world. In general, it is true that children do not work in factories, but it should not be totally inconceivable.

```
For scene SC1.2 to be built, the following conditions must be satisfied.
 A. The located object must be able to serve the function required by the reference object
       o This condition is not verified because a boy cannot work at a factory
 B. The situation must be normal
       * The size check is satisfied.
       * The materiality check is satisfied.
Scene SC1.2 cannot be built because all the conditions are not satisfied.
Conceptual representation C1.2 cannot be instantiated.

***** TRANSLATION INTO FRENCH *****

The description of scene SC1.1 follows:
'Object at landmark-in-highlighted-medium'
Both English and French use conceptual representation C1.1 to express this scene.
In addition, the context should be the following:
the-located-object-must-be-on-a-defined-trajectory
The French translation of the sentence follows:
 : Le garcon est a l' usine
```

6.3.1.2 Reference object is a location with a purpose

This section gives two examples of case 4.2.1.5. English and French use the same conceptualization in sentences of the type described here. The first example involves a 'generic place': the seaside.

```
 : The boy is at the seaside

***** INITIALIZATION *****

The located and the reference objects are not lexically ambiguous.

In English, 'seaside', when used as a reference object, can be conceptualized as:
 C1. a POINT if its highlighted property is one of the following:
    1. linear-object-focused-in-on-one-point
    2. location-with-a-purpose
 C2. a LINE if if its highlighted property is one of the following:
    1. fixed-line-of-reference
 C3.  SURFACE if its highlighted property is one of the following:
    1. part-of-object-support
    2. uncovered-geographic-entity


**** DERIVATION OF THE OBJECTIVE MEANING ****

Given the ideal meanings of 'at', only set C1 is selected.

For scene SC1.1 to be built, the following conditions must be satisfied.
 A. The located  object must be fix and large
       o These conditions are not verified because
        - a boy is not fixed or not large-sized.
 B. The situation must be normal
       * The size check is satisfied.
       * The materiality check is satisfied.
```

```
Scene SC1.1 cannot be built because all the conditions are not satisfied.
Conceptual representation C1.1 cannot be instantiated.

For scene SC1.2 to be built, the following conditions must be satisfied.
 A. The located object must be able to serve the function required by the reference object
       * This condition is verified because a boy can take-vacations at a seaside
 B. The situation must be normal
       * The size check is satisfied.
       * The materiality check is satisfied.
Because all the conditions are verified, scene SC1.2 is built.

***** TRANSLATION INTO FRENCH *****

The description of scene SC1.2 follows:
'Object at location-with-a-purpose'
Both English and French use conceptual representation C1.2 to express this scene.
The French translation of the sentence follows:
 : Le garcon est au  bord-de-la-mer

 CPU time : 93.16     GC time : 0.00
```

The second example involves a 'functional place': the supermarket. The second example is conceptually ambiguous, as mentioned in section 6.3.1.1, and belongs to the case of section 6.3.1.1 and the current one simultaneously. Note that the ambiguity detected in the second example is simple conceptual ambiguity, since the sentence is translated into a single ambiguous French sentence (however, the system cannot distinguish simple from complex conceptual ambiguity, and in the case of simple ambiguity, the French sentence is shown twice).

```
 : The boy is at the supermarket

***** INITIALIZATION *****

The located and the reference objects are not lexically ambiguous.

In English, 'supermarket', when used as a reference object, can be conceptualized as:
 C1. a POINT if its highlighted property is one of the following:
    1. landmark-in-highlighted-medium
    2. location-with-a-purpose
 C2.  SURFACE if its highlighted property is one of the following:
    1. plane-horizontal-support
    2. part-of-object-support
 C3. an EMPTY VOLUME if its highlighted property is one of the following:
    1. well-defined-low-container


**** DERIVATION OF THE OBJECTIVE MEANING ****

Given the ideal meanings of 'at', only set C1 is selected.

For scene SC1.1 to be built, the following conditions must be satisfied.
 A. The located object must be mobile
```

```
      * This condition is verified because a boy is mobile.
 B. The situation must be normal
      * The size check is satisfied.
      * The materiality check is satisfied.
Because all the conditions are verified, scene SC1.1 is built.

For scene SC1.2 to be built, the following conditions must be satisfied.
 A. The located object must be able to serve the function required by the reference object
      * This condition is verified because a boy can shop at a supermarket
 B. The situation must be normal
      * The size check is satisfied.
      * The materiality check is satisfied.
Because all the conditions are verified, scene SC1.2 is built.

***** TRANSLATION INTO FRENCH *****

*********************************
* The sentence is conceptually *
* ambiguous                    *
*********************************


The description of scene SC1.1 follows:
'Object at landmark-in-highlighted-medium'
Both English and French use conceptual representation C1.1 to express this scene.
In addition, the context should be the following:
the-located-object-must-be-on-a-defined-trajectory
The French translation of the sentence follows:
 : Le garcon  est  au  supermarche

The description of scene SC1.2 follows:
'Object at location-with-a-purpose'
Both English and French use conceptual representation C1.2 to express this scene.
The French translation of the sentence follows:
 : Le garcon  est  au  supermarche

 CPU time : 108.63     GC time : 0.00
```

6.3.1.3 Reference object is an artifact with a given purpose

In this case, English and French again use the same conceptualization. Recall that the only differences between this case and the previous one (6.3.1.2) are that in this case, the reference object is of size smaller than (or equal to) the prototypical size of a human being and the located object must not be contained in the reference object. Larger sizes are required and containment allowed in the sentences of section 6.3.1.2.

```
 : The girl is at the door

***** INITIALIZATION *****

The located and the reference objects are not lexically ambiguous.

In English, 'door', when used as a reference object, can be conceptualized as:
 C1. a POINT if its highlighted property is one of the following:
    1. landmark-in-highlighted-medium
    2. artifact-with-a-purpose
 C2.  SURFACE if its highlighted property is one of the following:
    1. plane-horizontal-support
    2. part-of-object-support


**** DERIVATION OF THE OBJECTIVE MEANING ****

Given the ideal meanings of 'at', only set C1 is selected.

For scene SC1.1 to be built, the following conditions must be satisfied.
 A. The located object must be mobile
       * This condition is verified because a girl is mobile.
 B. The situation must be normal
       * The size check is satisfied.
       * The materiality check is satisfied.
Because all the conditions are verified, scene SC1.1 is built.

For scene SC1.2 to be built, the following conditions must be satisfied.
 A. The located object must be able to serve the function required by the reference object
       * This condition is verified because a girl can enter at a door
 B. The situation must be normal
       * The size check is satisfied.
       * The materiality check is satisfied.
Because all the conditions are verified, scene SC1.2 is built.

***** TRANSLATION INTO FRENCH *****

*********************************
* The sentence is conceptually *
* ambiguous                     *
*********************************


The description of scene SC1.1 follows:
'Object at landmark-in-highlighted-medium'
Both English and French use conceptual representation C1.1 to express this scene.
In addition, the context should be the following:
the-located-object-must-be-on-a-defined-trajectory
The French translation of the sentence follows:
 : La fille est a la porte

The description of scene SC1.2 follows:
'Object at artifact-with-a-purpose'
Both English and French use conceptual representation C1.2 to express this scene.
The French translation of the sentence follows:
```

```
 : La fille est a la porte

 CPU time : 123.49     GC time : 0.00
```

6.3.1.4 Reference object is a generic sea

In this case, English and French use a different conceptualization.

```
 : The boys are at sea

***** INITIALIZATION *****

The located and the reference objects are not lexically ambiguous.

In English, 'sea', when used as a reference object, can be conceptualized as:
 C1. a POINT if its highlighted property is one of the following:
    1. generic-sea
 C2. a LINE if if its highlighted property is one of the following:
    1. body-of-water-whose-edge-is-focused-on
 C3.  SURFACE if its highlighted property is one of the following:
    1. part-of-object-support
    2. uncovered-geographic-entity
 C4. a FULL VOLUME if its highlighted property is one of the following:
    1. object-apprehended-as-a-whole


**** DERIVATION OF THE OBJECTIVE MEANING ****

Given the ideal meanings of 'at', only set C1 is selected.

For scene SC1.1 to be built, the following conditions must be satisfied.
 A. The located object must be smaller than a house
       * This condition is verified because a boy is smaller than a house.
 B. The situation must be normal
       * The size check is satisfied.
       * The materiality check is satisfied.
Because all the conditions are verified, scene SC1.1 is built.

***** TRANSLATION INTO FRENCH *****

The description of scene SC1.1 follows:
'Object at generic-sea'
The scene is expressed with different conceptualizations in English and French
The French conceptualization is: functional-container
The French translation of the sentence follows:
 : Les garcons sont en  mer

 CPU time : 98.57     GC time : 0.00
```

6.3.2 Conceptualizations as a line

This section shows runs of sentences that belong to cases described in section 4.2.2 of chapter 4.

6.3.2.1 Reference object is a body of water whose edge is focused on

In this example, the French conceptualization is different in English and in French. In French, the conceptualization used is 'water body edge', as indicated in the output. As well, this example illustrates a case of simple lexical ambiguity: the reference object has two meanings in French: 'rivière' (small river) and 'fleuve' (large river). Note that for lexical ambiguity, the system is able to indicate whether the sentence presents a case of simple or complex ambiguity. This ambiguity, since it is simple, does not cause a different conceptualization to be used in the spatial relation for each different French meaning of the ambiguous noun.

Another remark is that in this case, the system uses the two ideal meanings of the preposition 'on': 'Object contiguous with or supported by a LINE' and 'Object contiguous with or supported by a SURFACE'. In the case of the example of this section, however, only the conceptualization of the first is found to have been used in the sentence.

Another remark about the example of this section also concerns the example of the next section (6.3.2.2). In each of these two examples, French uses a different conceptualization, according to the fact that the edge of a water body, or that of a solid entity are considered.

```
 : The house on the river

***** INITIALIZATION *****

The reference objects is lexically ambiguous.
It is a case of simple lexical ambiguity.
************************************************
* Meaning 1 of the located object              *
* and meaning 1 and 2 of the reference object  *
* will now be processed                        *
************************************************


In English, 'river', when used as a reference object, can be conceptualized as:
 C1. a POINT if its highlighted property is one of the following:
    1. linear-object-focused-in-on-one-point
 C2. a LINE if if its highlighted property is one of the following:
    1. body-of-water-whose-edge-is-focused-on
 C3.  SURFACE if its highlighted property is one of the following:
    1. part-of-object-support
    2. uncovered-geographic-entity
 C4. a FULL VOLUME if its highlighted property is one of the following:
    1. object-apprehended-as-a-whole
 C5. an EMPTY VOLUME if its highlighted property is one of the following:
    1. well-defined-low-container


**** DERIVATION OF THE OBJECTIVE MEANING ****

Given the ideal meanings of 'on', only sets C2 and C3 are selected.
```

```
For scene SC3.1 to be built, the following conditions must be satisfied.
 A. The located object must be a part of the reference object
       o This condition is not verified because a river is not part of a house
 B. The situation must be normal
       * The size of the objects does not matter.
       * The materiality check is satisfied.
Scene SC3.1 cannot be built because all the conditions are not satisfied.
Conceptual representation C3.1 cannot be instantiated.

For scene SC3.2 to be built, the following conditions must be satisfied.
 A. The located object must be atmospheric
       o This condition is not verified because a house is not atmospheric.
 B. The situation must be normal
       * The size check is satisfied.
       * The materiality check is satisfied.
Scene SC3.2 cannot be built because all the conditions are not satisfied.
Conceptual representation C3.2 cannot be instantiated.

For scene SC2.1 to be built, the following conditions must be satisfied.
 A. The located  object must be fix and large
       * These conditions are verified because
        - a house is fixed and large-sized
 B. The situation must be normal
       * The size check is satisfied.
       * The materiality check is satisfied.
Because all the conditions are verified, scene SC2.1 is built.

***** TRANSLATION INTO FRENCH *****

The description of scene SC2.1 follows:
'Object on body-of-water-whose-edge-is-focused-on'
The scene is expressed with different conceptualizations in English and French
The French conceptualization is: water-body-edge
The French translation of the sentence follows:
 : La maison au bord de la riviere

The description of scene SC2.1 follows:
'Object on body-of-water-whose-edge-is-focused-on'
The scene is expressed with different conceptualizations in English and French
The French conceptualization is: water-body-edge
The French translation of the sentence follows:
 : La maison au bord du  fleuve

 CPU time : 112.10     GC time : 0.00
```

6.3.2.2 Reference object is a solid geographic entity

In this case, English and French use a different conceptualization. Also, note the difference of conceptualization in French in this case and that of section 6.3.2.1.

```
 : The house on the park

***** INITIALIZATION *****

The located and the reference objects are not lexically ambiguous.

In English, 'park', when used as a reference object, can be conceptualized as:
 C1. a POINT if its highlighted property is one of the following:
    1. landmark-in-highlighted-medium
 C2. a LINE if if its highlighted property is one of the following:
    1. solid-geographic-entity-whose-edge-is-focused-on
 C3.  SURFACE if its highlighted property is one of the following:
    1. part-of-object-support
    2. uncovered-geographic-entity
 C4. an EMPTY VOLUME if its highlighted property is one of the following:
    1. well-defined-low-container


**** DERIVATION OF THE OBJECTIVE MEANING ****

Given the ideal meanings of 'on', only sets C2 and C3 are selected.

For scene SC3.1 to be built, the following conditions must be satisfied.
 A. The located object must be a part of the reference object
       o This condition is not verified because a park is not part of a house
 B. The situation must be normal
       * The size of the objects does not matter.
       * The materiality check is satisfied.
Scene SC3.1 cannot be built because all the conditions are not satisfied.
Conceptual representation C3.1 cannot be instantiated.

For scene SC3.2 to be built, the following conditions must be satisfied.
 A. The located object must be atmospheric
       o This condition is not verified because a house is not atmospheric.
 B. The situation must be normal
       * The size check is satisfied.
       * The materiality check is satisfied.
Scene SC3.2 cannot be built because all the conditions are not satisfied.
Conceptual representation C3.2 cannot be instantiated.

For scene SC2.1 to be built, the following conditions must be satisfied.
 A. The located  object must be fix and large
       * These conditions are verified because
        - a house is fixed and large-sized
 B. The situation must be normal
       * The size check is satisfied.
       * The materiality check is satisfied.
 Because all the conditions are verified, scene SC2.1 is built.

***** TRANSLATION INTO FRENCH *****

The description of scene SC2.1 follows:
'Object on solid-geographic-entity-whose-edge-is-focused-on'
The scene is expressed with different conceptualizations in English and French
The French conceptualization is: open-space-edge
```

```
The French translation of the sentence follows:
 : La maison en bordure du  parc

 CPU time : 136.16     GC time : 0.00
```

6.3.3 Conceptualizations as a surface

This section shows runs of sentences that belong to cases described in section 4.2.3 of chapter 4.

6.3.3.1 Reference object is a plane horizontal support

In this case, English and French use the same conceptualization.

```
 : The boy is on the bed

***** INITIALIZATION *****

The located and the reference objects are not lexically ambiguous.

In English, 'bed', when used as a reference object, can be conceptualized as:
 C1. a POINT if its highlighted property is one of the following:
    1. landmark-in-highlighted-medium
 C2.  SURFACE if its highlighted property is one of the following:
    1. plane-horizontal-support
    2. part-of-object-support
 C3. an EMPTY VOLUME if its highlighted property is one of the following:
    1. well-defined-high-container


**** DERIVATION OF THE OBJECTIVE MEANING ****

Given the ideal meanings of 'on', only set C2 is selected.

For scene SC2.1 to be built, the following conditions must be satisfied.
 A. The reference object must be able to support the located object
       * This condition is verified because a bed can support a boy
 B. The situation must be normal
       * The size check is satisfied.
       * The materiality check is satisfied.
Because all the conditions are verified, scene SC2.1 is built.

For scene SC2.2 to be built, the following conditions must be satisfied.
 A. The located object must be a part of the reference object
       o This condition is not verified because a bed is not part of a boy
 B. The situation must be normal
       * The size of the objects does not matter.
       * The materiality check is satisfied.
Scene SC2.2 cannot be built because all the conditions are not satisfied.
Conceptual representation C2.2 cannot be instantiated.
```

```
***** TRANSLATION INTO FRENCH *****

The description of scene SC2.1 follows:
'Object on plane-horizontal-support'
Both English and French use conceptual representation C2.1 to express this scene.
The French translation of the sentence follows:
 : Le garcon est sur le lit

 CPU time : 170.71     GC time : 0.00
```

6.3.3.2 Reference object is a part of another which it can support

In this case, English and French use the same conceptualization.

```
 : The boy on the head

***** INITIALIZATION *****

The located and the reference objects are not lexically ambiguous.

In English, 'head', when used as a reference object, can be conceptualized as:
 C1.  SURFACE if its highlighted property is one of the following:
    1. plane-horizontal-support
    2. part-of-object-support
 C2. a FULL VOLUME if its highlighted property is one of the following:
    1. object-apprehended-as-a-whole


**** DERIVATION OF THE OBJECTIVE MEANING ****

Given the ideal meanings of 'on', only set C1 is selected.

For scene SC1.1 to be built, the following conditions must be satisfied.
 A. The reference object must be able to support the located object
       o This condition is not verified because a head cannot support a boy
 B. The situation must be normal
       * The size check is satisfied.
       * The materiality check is satisfied.
Scene SC1.1 cannot be built because all the conditions are not satisfied.
Conceptual representation C1.1 cannot be instantiated.

For scene SC1.2 to be built, the following conditions must be satisfied.
 A. The reference object must be a part of the located object
       * This condition is verified because a head is part of a boy
 B. The situation must be normal
       * The size of the objects does not matter.
       * The materiality check is satisfied.
Because all the conditions are verified, scene SC1.2 is built.

***** TRANSLATION INTO FRENCH *****

The description of scene SC1.2 follows:
```

```
'Object on part-of-object-support'
Both English and French use conceptual representation C1.2 to express this scene.
The French translation of the sentence follows:
 : Le garcon sur la tete

 CPU time : 93.79     GC time : 0.00
```

6.3.3.3 Non-horizontal object is involved in an intermediate support relation and is a wall

The example of this section, "The picture is on the wall", is very interesting for two different reasons. First of all, it is a case where two types of ambiguities occur at once: lexical ambiguity (one case of simple lexical ambiguity) and conceptual ambiguity (two cases of conceptual ambiguity). The case of lexical ambiguity involves the located object: 'picture' has two meanings in French: 'photo' (photograph) and 'tableau' (painting). Recall that since it is a case of simple lexical ambiguity, the use of one noun over the other does not affect the expression of the spatial relation. There are two conceptual ambiguities within the sentence. The first one is due to the fact that a 'picture' can be involved in an *intermediate* support relationship (see sections 4.2.3.5 and 4.2.3.6) and for this reason the French conceptualization in used is either SURFACE or POINT. The second one is due to the fact that a picture is fixed. For that reason, it can be considered as being 'on the wall', in the sense of 'against the wall', and the OPPOSING-FORCE conceptualization is used (this interpretation, however, is questionable).

The second reason that this example is very interesting involves the example of section 6.3.3.4. In the example of section 6.3.3.4, "The fly is on the wall", the 'fly' can be involved in an *active* support relation only. For this reason, only the SURFACE conceptualization can be used, whereas in the example currently considered, as mentioned above, both SURFACE and POINT conceptualizations are acceptable, since a 'picture' can be involved in an *intermediate* relation.

```
 : The picture is on the wall

***** INITIALIZATION *****

The located objects is lexically ambiguous.
It is a case of simple lexical ambiguity.
************************************************
* Meaning 1 and 2 of the located object        *
* and meaning 1 of the reference object        *
* will now be processed                        *
************************************************

In English, 'wall', when used as a reference object, can be conceptualized as:
 C1. a POINT if its highlighted property is one of the following:
    1. landmark-in-highlighted-medium
```

```
 C2.  SURFACE if its highlighted property is one of the following:
    1. part-of-object-support
    2. non-horizontal-object-involved-in-active-support-relation
    3. non-horizontal-object-involved-in-passive-support-relation
    4. wall-that-can-be-contiguous-with-other-objects
 C3. a FULL VOLUME if its highlighted property is one of the following:
    1. object-apprehended-as-a-whole


**** DERIVATION OF THE OBJECTIVE MEANING ****

Given the ideal meanings of 'on', only set C2 is selected.

For scene SC2.1 to be built, the following conditions must be satisfied.
 A. The located object must be a part of the reference object
       o This condition is not verified because a wall is not part of a picture
 B. The situation must be normal
       * The size of the objects does not matter.
       * The materiality check is satisfied.
Scene SC2.1 cannot be built because all the conditions are not satisfied.
Conceptual representation C2.1 cannot be instantiated.

For scene SC2.2 to be built, the following conditions must be satisfied.
 A. The located object must contribute to the adherence relation
       * This condition is verified because a picture is either adherent or can
         contribute to an adherence relation.
 B. The situation must be normal
       * The size check is satisfied.
       * The materiality check is satisfied.
Because all the conditions are verified, scene SC2.2 is built.

For scene SC2.3 to be built, the following conditions must be satisfied.
 A. The located object must not contribute to the adherence relation
       * This condition is verified because a picture is either non-adherent or
         can contribute to an adherence relation.
 B. The situation must be normal
       * The size check is satisfied.
       * The materiality check is satisfied.
Because all the conditions are verified, scene SC2.3 is built.

For scene SC2.4 to be built, the following conditions must be satisfied.
 A. The located object must be fix
       * This condition is verified because a picture is immobile.
 B. The situation must be normal
       * The size check is satisfied.
       * The materiality check is satisfied.
Because all the conditions are verified, scene SC2.4 is built.

***** TRANSLATION INTO FRENCH *****

*********************************
* The sentence is conceptually *
* ambiguous                     *
*********************************
```

```
The description of scene SC2.2 follows:
'Object on non-horizontal-object-involved-in-active-support-relation'
Both English and French use conceptual representation C2.2 to express this scene.
The French translation of the sentence follows:
 : La photo  est  sur le mur

The description of scene SC2.3 follows:
'Object on non-horizontal-object-involved-in-passive-support-relation'
The scene is expressed with different conceptualizations in English and French
The French conceptualization is: point
The French translation of the sentence follows:
 : La photo  est  au  mur

The description of scene SC2.4 follows:
'Object on wall-that-can-be-contiguous-with-other-objects'
The scene is expressed with different conceptualizations in English and French
The French conceptualization is: opposing-force
The French translation of the sentence follows:
 : La photo  est  contre le mur

The description of scene SC2.2 follows:
'Object on non-horizontal-object-involved-in-active-support-relation'
Both English and French use conceptual representation C2.4 to express this scene.
The French translation of the sentence follows:
 : Le tableau  est  sur le mur

The description of scene SC2.3 follows:
'Object on non-horizontal-object-involved-in-passive-support-relation'
The scene is expressed with different conceptualizations in English and French
The French conceptualization is: point
The French translation of the sentence follows:
 : Le tableau  est  au  mur

The description of scene SC2.4 follows:
'Object on wall-that-can-be-contiguous-with-other-objects'
The scene is expressed with different conceptualizations in English and French
The French conceptualization is: opposing-force
The French translation of the sentence follows:
 : Le tableau  est  contre le mur

 CPU time : 294.10     GC time : 0.00
```

6.3.3.4 Non-horizontal object is involved in a passive support relation

In this case, English and French use the same conceptualization. Refer to section 6.3.3.3 for comparison.

```
 : The fly is on the wall

***** INITIALIZATION *****

The located and the reference objects are not lexically ambiguous.

In English, 'wall', when used as a reference object, can be conceptualized as:
 C1. a POINT if its highlighted property is one of the following:
    1. landmark-in-highlighted-medium
 C2.  SURFACE if its highlighted property is one of the following:
    1. part-of-object-support
    2. non-horizontal-object-involved-in-active-support-relation
    3. non-horizontal-object-involved-in-passive-support-relation
    4. wall-that-can-be-contiguous-with-other-objects
 C3. a FULL VOLUME if its highlighted property is one of the following:
    1. object-apprehended-as-a-whole


**** DERIVATION OF THE OBJECTIVE MEANING ****

Given the ideal meanings of 'on', only set C2 is selected.

For scene SC2.1 to be built, the following conditions must be satisfied.
 A. The located object must be a part of the reference object
       o This condition is not verified because a wall is not part of a fly
 B. The situation must be normal
       * The size of the objects does not matter.
       * The materiality check is satisfied.
Scene SC2.1 cannot be built because all the conditions are not satisfied.
Conceptual representation C2.1 cannot be instantiated.

For scene SC2.2 to be built, the following conditions must be satisfied.
 A. The located object must contribute to the adherence relation
       * This condition is verified because a fly is either adherent or can
         contribute to an adherence relation.
 B. The situation must be normal
       * The size check is satisfied.
       * The materiality check is satisfied.
Because all the conditions are verified, scene SC2.2 is built.

For scene SC2.3 to be built, the following conditions must be satisfied.
 A. The located object must not contribute to the adherence relation
       o This condition is not verified because a fly is adherent.
 B. The situation must be normal
       * The size check is satisfied.
       * The materiality check is satisfied.
Scene SC2.3 cannot be built because all the conditions are not satisfied.
Conceptual representation C2.3 cannot be instantiated.

For scene SC2.4 to be built, the following conditions must be satisfied.
 A. The located object must be fix
       o This condition is not verified because a fly is not immobile.
 B. The situation must be normal
       * The size check is satisfied.
       * The materiality check is satisfied.
```

```
Scene SC2.4 cannot be built because all the conditions are not satisfied.
Conceptual representation C2.4 cannot be instantiated.

***** TRANSLATION INTO FRENCH *****

The description of scene SC2.2 follows:
'Object on non-horizontal-object-involved-in-active-support-relation'
Both English and French use conceptual representation C2.2 to express this scene.
The French translation of the sentence follows:
 : La mouche  est  sur le mur

 CPU time : 212.80     GC time ; 0.00
```

6.3.3.5 Reference object is an horizontal support and has sides

This section shows another set of interesting examples. The first example shows cases of simple lexical and complex lexical ambiguity and the second example shows a case of complex conceptual ambiguity.

In the first example,"The cat is on the street", the reference object has a complex lexical ambiguity, which leads to a different conceptualization of the relation according to the French meaning chosen. 'Street' has two meanings in French: 'rue', which takes into account the roadway, the sidewalks, and the houses surrounding the street (roadway), so its conceptualization results in EMPTY VOLUME; and 'chaussée', which takes into consideration only the roadway, so its conceptualization results in SURFACE. These two different meanings are each represented by a different concept in the knowledge base: street and street2, respectively. Additionally, 'cat' can be translated as 'chat' (he-cat) or 'chatte' (she-cat), but this ambiguity is simple. The system therefore outputs four different sentences using different nouns and/or different prepositions. Note also that conceptual representation C3.3 has two highlights. This is because a street can both be seen as a supporting platform and a container.

```
 : The cat is on the street

***** INITIALIZATION *****

The located and the reference objects are lexically ambiguous.
It is a case of simple lexical ambiguity for the located object.
It is a case of complex lexical ambiguity for the reference object.
************************************************
* Meaning 1 and 2 of the located object        *
* and meaning 1 of the reference object        *
* will now be processed                        *
************************************************

In English, 'street', when used as a reference object, can be conceptualized as:
 C1. a POINT if its highlighted property is one of the following:
    1. linear-object-focused-in-on-one-point
```

```
 C2. a LINE if if its highlighted property is one of the following:
    1. fixed-line-of-reference
 C3.  SURFACE if its highlighted property is one of the following:
    1. part-of-object-support
    2. uncovered-geographic-entity
    3. object-with-horizontal-support
       object-with-sides

**** DERIVATION OF THE OBJECTIVE MEANING ****

Given the ideal meanings of 'on', only sets C2 and C3 are selected.

For scene SC3.1 to be built, the following conditions must be satisfied.
 A. The located object must be a part of the reference object
       o This condition is not verified because a street is not part of a cat
 B. The situation must be normal
       * The size of the objects does not matter.
       * The materiality check is satisfied.
Scene SC3.1 cannot be built because all the conditions are not satisfied.
Conceptual representation C3.1 cannot be instantiated.

For scene SC3.2 to be built, the following conditions must be satisfied.
 A. The located object must be atmospheric
       o This condition is not verified because a cat is not atmospheric.
 B. The situation must be normal
       * The size check is satisfied.
       * The materiality check is satisfied.
Scene SC3.2 cannot be built because all the conditions are not satisfied.
Conceptual representation C3.2 cannot be instantiated.

For scene SC3.3 to be built, the following conditions must be satisfied.
 A. The situation must be normal
       * The size check is satisfied.
       * The materiality check is satisfied.
 ---- No condition involving the located object needs to be checked ----
Because all the conditions are verified, scene SC3.3 is built.

For scene SC2.1 to be built, the following conditions must be satisfied.
 A. The located  object must be fix and large
       o These conditions are not verified because
        - a cat is not fixed or not large-sized.
 B. The situation must be normal
       * The size check is satisfied.
       * The materiality check is satisfied.
Scene SC2.1 cannot be built because all the conditions are not satisfied.
Conceptual representation C2.1 cannot be instantiated.

***** TRANSLATION INTO FRENCH *****

The description of scene SC3.3 follows:
'Object on object-with-horizontal-support'
The scene is expressed with different conceptualizations in English and French
The French conceptualization is: volume
The French translation of the sentence follows:
```

```
 : Le chat est dans la rue

The description of scene SC3.3 follows:
'Object on object-with-horizontal-support'
The scene is expressed with different conceptualizations in English and French
The French conceptualization is: volume
The French translation of the sentence follows:
 : La chatte est dans la rue

************************************************
* Meaning 1 and 2 of the located object        *
* and meaning 2 of the reference object        *
* will now be processed                        *
************************************************


In English, 'street', when used as a reference object, can be conceptualized as:
 C1. a POINT if its highlighted property is one of the following:
    1. linear-object-focused-in-on-one-point
 C2. a LINE if if its highlighted property is one of the following:
    1. fixed-line-of-reference
 C3.  SURFACE if its highlighted property is one of the following:
    1. plane-horizontal-support
    2. part-of-object-support
    3. uncovered-geographic-entity
 C4. an EMPTY VOLUME if its highlighted property is one of the following:
    1. through-way-that-can-be-obstructed


**** DERIVATION OF THE OBJECTIVE MEANING ****

Given the ideal meanings of 'on', only sets C2 and C3 are selected.

For scene SC3.1 to be built, the following conditions must be satisfied.
 A. The reference object must be able to support the located object
       * This condition is verified because a street2 can support a cat
 B. The situation must be normal
       * The size check is satisfied.
       * The materiality check is satisfied.
Because all the conditions are verified, scene SC3.1 is built.

For scene SC3.2 to be built, the following conditions must be satisfied.
 A. The located object must be a part of the reference object
       o This condition is not verified because a street2 is not part of a cat
 B. The situation must be normal
       * The size of the objects does not matter.
       * The materiality check is satisfied.
Scene SC3.2 cannot be built because all the conditions are not satisfied.
Conceptual representation C3.2 cannot be instantiated.

For scene SC3.3 to be built, the following conditions must be satisfied.
 A. The located object must be atmospheric
       o This condition is not verified because a cat is not atmospheric.
 B. The situation must be normal
       * The size check is satisfied.
```

```
       * The materiality check is satisfied.
Scene SC3.3 cannot be built because all the conditions are not satisfied.
Conceptual representation C3.3 cannot be instantiated.

For scene SC2.1 to be built, the following conditions must be satisfied.
 A. The located  object must be fix and large
       o These conditions are not verified because
        - a cat is not fixed or not large-sized.
 B. The situation must be normal
       * The size check is satisfied.
       * The materiality check is satisfied.
Scene SC2.1 cannot be built because all the conditions are not satisfied.
Conceptual representation C2.1 cannot be instantiated.

***** TRANSLATION INTO FRENCH *****

The description of scene SC3.1 follows:
'Object on plane-horizontal-support'
Both English and French use conceptual representation C3.1 to express this scene.
The French translation of the sentence follows:
 : Le chat est sur la chaussee

The description of scene SC3.1 follows:
'Object on plane-horizontal-support'
Both English and French use conceptual representation C3.1 to express this scene.
The French translation of the sentence follows:
 : La chatte est sur la chaussee

 CPU time : 265.39     GC time : 0.00
```

The second example, "The boy is on the bus", presents a case of complex conceptual ambiguity, in which the ambiguous English sentence is translated into two unambiguous French ones. In this case, the nouns are the same in the two translations, it is the preposition which is different. As mentioned previously, however, the system is not able to indicate whether the conceptual complexity is a case of simple or complex ambiguity. Note also that conceptual representation C2.3 has two highlights. This is because a bus can both be seen as a supporting platform and a container.

```
 : The boy is on the bus

***** INITIALIZATION *****

The located and the reference objects are not lexically ambiguous.

In English, 'bus', when used as a reference object, can be conceptualized as:
 C1. a POINT if its highlighted property is one of the following:
    1. landmark-in-highlighted-medium
 C2.  SURFACE if its highlighted property is one of the following:
    1. plane-horizontal-support
    2. part-of-object-support
    3. object-with-horizontal-support
       object-with-sides
 C3. a FULL VOLUME if its highlighted property is one of the following:
    1. object-apprehended-as-a-whole
```

```
 C4. an EMPTY VOLUME if its highlighted property is one of the following:
    1. well-defined-low-container

**** DERIVATION OF THE OBJECTIVE MEANING ****

Given the ideal meanings of 'on', only set C2 is selected.

For scene SC2.1 to be built, the following conditions must be satisfied.
 A. The reference object must be able to support the located object
       * This condition is verified because a bus can support a boy
 B. The situation must be normal
       * The size check is satisfied.
       * The materiality check is satisfied.
Because all the conditions are verified, scene SC2.1 is built.

For scene SC2.2 to be built, the following conditions must be satisfied.
 A. The located object must be a part of the reference object
       o This condition is not verified because a bus is not part of a boy
 B. The situation must be normal
       * The size of the objects does not matter.
       * The materiality check is satisfied.
Scene SC2.2 cannot be built because all the conditions are not satisfied.
Conceptual representation C2.2 cannot be instantiated.

For scene SC2.3 to be built, the following conditions must be satisfied.
 A. The situation must be normal
       * The size check is satisfied.
       * The materiality check is satisfied.
 --- No condition involving the located object needs to be checked ---
Because all the conditions are verified, scene SC2.3 is built.

***** TRANSLATION INTO FRENCH *****

*********************************
* The sentence is conceptually *
* ambiguous                    *
*********************************


The description of scene SC2.1 follows:
'Object on plane-horizontal-support'
Both English and French use conceptual representation C2.1 to express this scene.
The French translation of the sentence follows:
 : Le garcon  est  sur le bus

The description of scene SC2.3 follows:
'Object on object-with-horizontal-support'
The scene is expressed with different conceptualizations in English and French
The French conceptualization is: volume
The French translation of the sentence follows:
 : Le garcon  est  dans le bus

 CPU time : 183.83     GC time : 0.00
```

6.3.3.6 Reference object is an open-space support

In this case, English and French use the same conceptualization. We show two examples of this case: "Florence is on the continent" and "Norrin is on the square". The first example follows:

```
 : Florence is on the continent

***** INITIALIZATION *****

The located and the reference objects are not lexically ambiguous.

In English, 'continent', when used as a reference object, can be conceptualized as:
 C1. a LINE if if its highlighted property is one of the following:
    1. solid-geographic-entity-whose-edge-is-focused-on
 C2.  SURFACE if its highlighted property is one of the following:
    1. part-of-object-support
    2. uncovered-geographic-entity
    3. open-space-support


**** DERIVATION OF THE OBJECTIVE MEANING ****

Given the ideal meanings of 'on', only sets C1 and C2 are selected.

For scene SC2.1 to be built, the following conditions must be satisfied.
 A. The located object must be a part of the reference object
       o This condition is not verified because a continent is not part of a Florence
 B. The situation must be normal
       * The size of the objects does not matter.
       * The materiality check is satisfied.
Scene SC2.1 cannot be built because all the conditions are not satisfied.
Conceptual representation C2.1 cannot be instantiated.

For scene SC2.2 to be built, the following conditions must be satisfied.
 A. The located object must be atmospheric
       o This condition is not verified because a Florence is not atmospheric.
 B. The situation must be normal
       * The size check is satisfied.
       * The materiality check is satisfied.
Scene SC2.2 cannot be built because all the conditions are not satisfied.
Conceptual representation C2.2 cannot be instantiated.

For scene SC2.3 to be built, the following conditions must be satisfied.
 A. The situation must be normal
       * The size check is satisfied.
       * The materiality check is satisfied.
 --- No condition involving the located object needs to be checked ---
Because all the conditions are verified, scene SC2.3 is built.

For scene SC1.1 to be built, the following conditions must be satisfied.
 A. The located  object must be fix and large
       o These conditions are not verified because
        - a Florence is not fixed or not large-sized.
```

```
 B. The situation must be normal
       * The size check is satisfied.
       * The materiality check is satisfied.
Scene SC1.1 cannot be built because all the conditions are not satisfied.
Conceptual representation C1.1 cannot be instantiated.

***** TRANSLATION INTO FRENCH *****

The description of scene SC2.3 follows:
'Object on open-space-support'
Both English and French use conceptual representation C2.3 to express this scene.

 : Florence est sur le continent

 CPU time : 147.90     GC time : 0.00
```

The second example follows:

```
 : Norrin is on the square

***** INITIALIZATION *****

The located and the reference objects are not lexically ambiguous.

In English, 'square', when used as a reference object, can be conceptualized as:
 C1. a LINE if if its highlighted property is one of the following:
    1. solid-geographic-entity-whose-edge-is-focused-on
 C2.  SURFACE if its highlighted property is one of the following:
    1. part-of-object-support
    2. uncovered-geographic-entity
    3. open-space-support


**** DERIVATION OF THE OBJECTIVE MEANING ****

Given the ideal meanings of 'on', only sets C1 and C2 are selected.

For scene SC2.1 to be built, the following conditions must be satisfied.
 A. The located object must be a part of the reference object
       o This condition is not verified because a square is not part of a Norrin
 B. The situation must be normal
       * The size of the objects does not matter.
       * The materiality check is satisfied.
Scene SC2.1 cannot be built because all the conditions are not satisfied.
Conceptual representation C2.1 cannot be instantiated.

For scene SC2.2 to be built, the following conditions must be satisfied.
 A. The located object must be atmospheric
       o This condition is not verified because a Norrin is not atmospheric.
 B. The situation must be normal
       * The size check is satisfied.
       * The materiality check is satisfied.
Scene SC2.2 cannot be built because all the conditions are not satisfied.
Conceptual representation C2.2 cannot be instantiated.

For scene SC2.3 to be built, the following conditions must be satisfied.
```

```
 A. The situation must be normal
       * The size check is satisfied.
       * The materiality check is satisfied.
 --- No condition involving the located object needs to be checked ---
Because all the conditions are verified, scene SC2.3 is built.

For scene SC1.1 to be built, the following conditions must be satisfied.
 A. The located  object must be fix and large
       o These conditions are not verified because
        - a Norrin is not fixed or not large-sized.
 B. The situation must be normal
       * The size check is satisfied.
       * The materiality check is satisfied.
Scene SC1.1 cannot be built because all the conditions are not satisfied.
Conceptual representation C1.1 cannot be instantiated.

***** TRANSLATION INTO FRENCH *****

The description of scene SC2.3 follows:
'Object on open-space-support'
Both English and French use conceptual representation C2.3 to express this scene.
The French translation of the sentence follows:
 : Norrin est sur la place

 CPU time : 156.97     GC time : 0.00
```

6.3.4 Conceptualization as a bounded surface

This section shows sentences that belong to cases described in section 4.2.4 of chapter 4.

6.3.4.1 Reference object is a bounded surface and a country, a continent, or a city

The four examples of this section show interesting cases, outlining the French grammatical rules concerning cities, countries, and continent. We recall that these rules say that cities are conceptualized as POINTS, feminine names of countries and continents, or masculine ones starting with a vowel as FUNCTIONAL-CONTAINER, and masculine ones starting with a consonant as POINTS.[2]

It is interesting to compare examples of this section involving a country name with the example of section 6.3.4.2 and that involving a continent with the first example of section 6.3.3.6: when proper names are not used, the conceptualization of the same entity is different, and interestingly, that of countries differs from that of continents (countries and cities not referred to by their proper name have the same conceptualization).

Note that in the runs of this section, the SURFACE conceptualization is not selected

[2]Even though we treat these cases in the same way as the others, we acknowledge the fact that they are different, in that they seem purely conventional, rather than cognitive, as well. Future work on this issue is discussed in chapter 7.

because the BOUNDED-SURFACE conceptualization is, instead (see chapter 4).

In the case of this section, English and French use different conceptualizations in each instance.

We present four different examples of this case. The first example follows:

```
 : The girl is in Paris

***** INITIALIZATION *****

The located and the reference objects are not lexically ambiguous.

In English, 'Paris', when used as a reference object, can be conceptualized as:
 C1. a POINT if its highlighted property is one of the following:
    1. location-with-a-purpose
 C2.  SURFACE if its highlighted property is one of the following:
    1. part-of-object-support
    2. uncovered-geographic-entity
 C3. a BOUNDED-SURFACE if its highlighted property is one of the following:
    1. bounded-surface


**** DERIVATION OF THE OBJECTIVE MEANING ****

Given the ideal meanings of 'in', only set C3 is selected.

For scene SC3.1 to be built, the following conditions must be satisfied.
 A. The situation must be normal
       * The size check is satisfied.
       * The materiality check is satisfied.
 --- No condition involving the located object needs to be checked ---
Because all the conditions are verified, scene SC3.1 is built.

***** TRANSLATION INTO FRENCH *****

The description of scene SC3.1 follows:
'Object in bounded-surface'
The scene is expressed with different conceptualizations in English and French
The French conceptualization is: point
The French translation of the sentence follows:
 : La fille est a Paris

 CPU time : 138.85     GC time : 0.00
```

The second example follows:

```
 : Paris is in France

***** INITIALIZATION *****

The located and the reference objects are not lexically ambiguous.

In English, 'France', when used as a reference object, can be conceptualized as:
 C1.  SURFACE if its highlighted property is one of the following:
    1. part-of-object-support
    2. uncovered-geographic-entity
 C2. a BOUNDED-SURFACE if its highlighted property is one of the following:
    1. bounded-surface


**** DERIVATION OF THE OBJECTIVE MEANING ****

Given the ideal meanings of 'in', only set C2 is selected.

For scene SC2.1 to be built, the following conditions must be satisfied.
 A. The situation must be normal
       * The size check is satisfied.
       * The materiality check is satisfied.
 --- No condition involving the located object needs to be checked ---
Because all the conditions are verified, scene SC2.1 is built.

***** TRANSLATION INTO FRENCH *****

The description of scene SC2.1 follows:
'Object in bounded-surface'
The scene is expressed with different conceptualizations in English and French
The French conceptualization is: functional-container
The French translation of the sentence follows:
 : Paris est en France

 CPU time : 168.49     GC time : 0.00
```

The third example follows:

```
 : Norrin is in Europe

***** INITIALIZATION *****

The located and the reference objects are not lexically ambiguous.

In English, 'Europe', when used as a reference object, can be conceptualized as:
 C1.  SURFACE if its highlighted property is one of the following:
    1. part-of-object-support
    2. uncovered-geographic-entity
 C2. a BOUNDED-SURFACE if its highlighted property is one of the following:
    1. bounded-surface


**** DERIVATION OF THE OBJECTIVE MEANING ****

Given the ideal meanings of 'in', only set C2 is selected.
```

```
For scene SC2.1 to be built, the following conditions must be satisfied.
 A. The situation must be normal
       * The size check is satisfied.
       * The materiality check is satisfied.
 --- No condition involving the located object needs to be checked ---
Because all the conditions are verified, scene SC2.1 is built.

***** TRANSLATION INTO FRENCH *****

The description of scene SC2.1 follows:
'Object in bounded-surface'
The scene is expressed with different conceptualizations in English and French
The French conceptualization is: functional-container
The French translation of the sentence follows:
 : Norrin est en Europe

 CPU time : 71.12     GC time : 0.00
```

The fourth example follows:

```
 : Maxime is in Canada

***** INITIALIZATION *****

The located and the reference objects are not lexically ambiguous.

In English, 'Canada', when used as a reference object, can be conceptualized as:
 C1.  SURFACE if its highlighted property is one of the following:
    1. part-of-object-support
    2. uncovered-geographic-entity
 C2. a BOUNDED-SURFACE if its highlighted property is one of the following:
    1. bounded-surface


**** DERIVATION OF THE OBJECTIVE MEANING ****

Given the ideal meanings of 'in', only set C2 is selected.

For scene SC2.1 to be built, the following conditions must be satisfied.
 A. The situation must be normal
       * The size check is satisfied.
       * The materiality check is satisfied.
 --- No condition involving the located object needs to be checked ---
Because all the conditions are verified, scene SC2.1 is built.

***** TRANSLATION INTO FRENCH *****

The description of scene SC2.1 follows:
'Object in bounded-surface'
The scene is expressed with different conceptualizations in English and French
The French conceptualization is: point
The French translation of the sentence follows:
 : Maxime est au Canada
```

6.3.4.2 Reference object is a bounded surface

In this case English and French use the same conceptualization. Contrast with the previous section (6.3.4.1).

```
 : The boy is in the city

***** INITIALIZATION *****

The located and the reference objects are not lexically ambiguous.

In English, 'city', when used as a reference object, can be conceptualized as:
 C1. a POINT if its highlighted property is one of the following:
    1. location-with-a-purpose
 C2.  SURFACE if its highlighted property is one of the following:
    1. part-of-object-support
    2. uncovered-geographic-entity
 C3. a BOUNDED-SURFACE if its highlighted property is one of the following:
    1. bounded-surface


**** DERIVATION OF THE OBJECTIVE MEANING ****

Given the ideal meanings of 'in', only set C3 is selected.

For scene SC3.1 to be built, the following conditions must be satisfied.
 A. The situation must be normal
       * The size check is satisfied.
       * The materiality check is satisfied.
 --- No condition involving the located object needs to be checked ---
Because all the conditions are verified, scene SC3.1 is built.

***** TRANSLATION INTO FRENCH *****

The description of scene SC3.1 follows:
'Object in bounded-surface'
Both English and French use conceptual representation C3.1 to express this scene.
The French translation of the sentence follows:
 : Le garcon est dans la ville

 CPU time : 158.33     GC time : 0.00
```

6.3.5 Conceptualizations as a full volume

This section shows runs of sentences that belong to cases described in section 4.2.5 of chapter 4.

6.3.5.1 Reference object is apprehended as a whole

In this case English and French use the same conceptualization.

```
 : The hole in the shirt

***** INITIALIZATION *****

The located and the reference objects are not lexically ambiguous.

In English, 'shirt', when used as a reference object, can be conceptualized as:
 C1.  SURFACE if its highlighted property is one of the following:
    1. plane-horizontal-support
    2. part-of-object-support
 C2. a FULL VOLUME if its highlighted property is one of the following:
    1. object-apprehended-as-a-whole
 C3. an EMPTY VOLUME if its highlighted property is one of the following:
    1. piece-of-clothing


**** DERIVATION OF THE OBJECTIVE MEANING ****

Given the ideal meanings of 'in', only sets C2 and C3 are selected.

For scene SC2.1 to be built, the following conditions must be satisfied.
 A. The located object can be a constituent of the reference object
       * This condition is verified because a hole is a constituent
         of a shirt
 B. The situation must be normal
       * The size of the objects does not matter.
       * The materiality of the located object does not matter.
Because all the conditions are verified, scene SC2.1 is built.

For scene SC3.1 to be built, the following conditions must be satisfied.
 A. The located object must be human
       o This condition is not verified because a hole is not a human being.
 B. The situation must be normal
       * The size check is satisfied.
       o The materiality check is not satisfied.
Scene SC3.1 cannot be built because all the conditions are not satisfied.
Conceptual representation C3.1 cannot be instantiated.

***** TRANSLATION INTO FRENCH *****

The description of scene SC2.1 follows:
'Object in object-apprehended-as-a-whole'
Both English and French use conceptual representation C2.1 to express this scene.
The French translation of the sentence follows:
 : Le trou dans la chemise

 CPU time : 84.64     GC time : 0.00
```

6.3.6 Conceptualizations as an empty volume

This section shows runs of sentences that belong to cases described in section 4.2.6 of chapter 4.

6.3.6.1 Reference object is a well defined low container

In this case English and French use the same conceptualization. A counterpart of the example of this section, in which the hermeticity condition involving the located object, is not verified and the sentence invalid, is given in section 6.4.

```
 : The water is in the glass

***** INITIALIZATION *****

The located and the reference objects are not lexically ambiguous.

In English, 'glass', when used as a reference object, can be conceptualized as:
 C1.  SURFACE if its highlighted property is one of the following:
    1. part-of-object-support
 C2. a FULL VOLUME if its highlighted property is one of the following:
    1. object-apprehended-as-a-whole
 C3. an EMPTY VOLUME if its highlighted property is one of the following:
    1. well-defined-low-container


**** DERIVATION OF THE OBJECTIVE MEANING ****

Given the ideal meanings of 'in', only sets C2 and C3 are selected.

For scene SC2.1 to be built, the following conditions must be satisfied.
 A. The located object can be a constituent of the reference object
       o This condition is not verified because a water is not a constituent
         of a glass
 B. The situation must be normal
       * The size check is satisfied.
       * The materiality check is satisfied.
Scene SC2.1 cannot be built because all the conditions are not satisfied.
Conceptual representation C2.1 cannot be instantiated.

For scene SC3.1 to be built, the following conditions must be satisfied.
 A. The reference object must be hermetic to the located object
       * This condition is verified because
         - a water is a liquid
         - a glass is hermetic to liquids
 B. The situation must be normal
       * The size check is satisfied.
       * The materiality check is satisfied.
Because all the conditions are verified, scene SC3.1 is built.

***** TRANSLATION INTO FRENCH *****

The description of scene SC3.1 follows:
'Object in well-defined-low-container'
Both English and French use conceptual representation C3.1 to express this scene.
The French translation of the sentence follows:
 : L' eau est dans le verre
```

6.3.6.2 Reference object is a well defined low container and not a well defined high one

The example here presents a simple case of conceptual ambiguity, in which the tree can be thought as 'the trunk of the tree' or 'the branches of the tree'. English and French do not distinguish between these two meanings. Also, they both use the same conceptualization: EMPTY VOLUME.

```
 : The boy is in the tree

***** INITIALIZATION *****

The located and the reference objects are not lexically ambiguous.

In English, 'tree', when used as a reference object, can be conceptualized as:
 C1. a POINT if its highlighted property is one of the following:
    1. landmark-in-highlighted-medium
 C2.  SURFACE if its highlighted property is one of the following:
    1. part-of-object-support
    2. non-horizontal-object-involved-in-active-support-relation
    3. non-horizontal-object-involved-in-passive-support-relation
 C3. a FULL VOLUME if its highlighted property is one of the following:
    1. object-apprehended-as-a-whole
 C4. an EMPTY VOLUME if its highlighted property is one of the following:
    1. well-defined-low-container
    2. non-well-defined-high-container


**** DERIVATION OF THE OBJECTIVE MEANING ****

Given the ideal meanings of 'in', only sets C3 and C4 are selected.

For scene SC3.1 to be built, the following conditions must be satisfied.
 A. The located object can be a constituent of the reference object
       o This condition is not verified because a boy is not a constituent
         of a tree
 B. The situation must be normal
       * The size check is satisfied.
       * The materiality check is satisfied.
Scene SC3.1 cannot be built because all the conditions are not satisfied.
Conceptual representation C3.1 cannot be instantiated.

For scene SC4.1 to be built, the following conditions must be satisfied.
 A. The reference object must be hermetic to the located object
       * This condition is verified because
         - a boy is a solid
         - a tree is hermetic to solids
 B. The situation must be normal
       * The size check is satisfied.
       * The materiality check is satisfied.
Because all the conditions are verified, scene SC4.1 is built.

For scene SC4.2 to be built, the following conditions must be satisfied.
 A. The reference object must be hermetic to the located object
```

```
        * This condition is verified because
          - a boy is a solid
          - a tree is hermetic to solids
 B. The situation must be normal
       * The size check is satisfied.
       * The materiality check is satisfied.
Because all the conditions are verified, scene SC4.2 is built.

***** TRANSLATION INTO FRENCH *****

*********************************
* The sentence is conceptually *
* ambiguous                    *
*********************************


The description of scene SC4.1 follows:
'Object in well-defined-low-container'
Both English and French use conceptual representation C4.1 to express this scene.
The French translation of the sentence follows:
 : Le garcon  est  dans l' arbre

The description of scene SC4.2 follows:
'Object in non-well-defined-high-container'
Both English and French use conceptual representation C4.2 to express this scene.
The French translation of the sentence follows:
 : Le garcon  est  dans l' arbre

 CPU time : 203.89     GC time : 0.00
```

6.3.6.3 Reference object is a part of space or environment and cannot hide

This case is interesting when compared with the cases of the next two sections (6.3.6.4 and 6.3.6.5). In each of these three cases, English uses the same conceptualization, whereas French uses a different one, distinguishing cases where the reference object has an almost nonexistent width (6.3.6.4) or is an environment in which one can hide (6.3.6.5). In the current case, the French conceptualization used is: POINT; in section 6.3.6.4, it is SURFACE; and in 6.3.6.5, it is EMPTY VOLUME.

Note that these three runs output similar information. This is because the output takes the English point of view, which considers the three cases similar. What is not shown is that in French, a different conceptualization is built in each of the cases that matches the English one (in each case, only the French conceptualization corresponding to the case is built: the other two are not).

Note that here again, C2.1 has two highlights.

```
 : Joshua is in the shade

***** INITIALIZATION *****

The located and the reference objects are not lexically ambiguous.

In English, 'shade', when used as a reference object, can be conceptualized as:
 C1.  SURFACE if its highlighted property is one of the following:
    1. part-of-object-support
 C2. an EMPTY VOLUME if its highlighted property is one of the following:
    1. environment-in-which-one-can-hide-cannot-hide-or-which-has-a-tiny-width
       part-of-space-or-environment


**** DERIVATION OF THE OBJECTIVE MEANING ****

Given the ideal meanings of 'in', only set C2 is selected.

For scene SC2.1 to be built, the following conditions must be satisfied.
 A. The situation must be normal
       * The size check is satisfied.
       * The materiality check is satisfied.
 --- No condition involving the located object needs to be checked ---
Because all the conditions are verified, scene SC2.1 is built.

***** TRANSLATION INTO FRENCH *****

The description of scene SC2.1 follows:
'Object in environment-in-which-one-can-hide-cannot-hide-or-which-has-a-tiny-width'
The scene is expressed with different conceptualizations in English and French
The French conceptualization is: point
The French translation of the sentence follows:
 : Joshua est a l' ombre

 CPU time : 60.60     GC time : 0.00
```

6.3.6.4 Reference object is a part of space or environment and has an almost inexistent width

In this case English and French use a different conceptualization. Compare with sections 6.3.6.3 and 6.3.6.5. As well, note that the sentence is lexically ambiguous.

```
 : Marc-Alain is in the picture

***** INITIALIZATION *****

The reference objects is lexically ambiguous.
It is a case of simple lexical ambiguity.
************************************************
* Meaning 1 of the located object              *
* and meaning 1 and 2 of the reference object  *
* will now be processed                        *
************************************************
```

```
In English, 'picture', when used as a reference object, can be conceptualized as:
 C1.  SURFACE if its highlighted property is one of the following:
    1. plane-horizontal-support
    2. part-of-object-support
 C2. an EMPTY VOLUME if its highlighted property is one of the following:
    1. environment-in-which-one-can-hide-cannot-hide-or-which-has-a-tiny-width
       part-of-space-or-environment


**** DERIVATION OF THE OBJECTIVE MEANING ****

Given the ideal meanings of 'in', only set C2 is selected.

For scene SC2.1 to be built, the following conditions must be satisfied.
 A. The situation must be normal
       * The size check is satisfied.
       * The materiality check is satisfied.
 --- No condition involving the located object needs to be checked ---
Because all the conditions are verified, scene SC2.1 is built.

***** TRANSLATION INTO FRENCH *****

The description of scene SC2.1 follows:
'Object in environment-in-which-one-can-hide-cannot-hide-or-which-has-a-tiny-width'
The scene is expressed with different conceptualizations in English and French
The French conceptualization is: surface
The French translation of the sentence follows:
 : Marc-Alain est sur la photo

The description of scene SC2.1 follows:
'Object in environment-in-which-one-can-hide-cannot-hide-or-which-has-a-tiny-width'
The scene is expressed with different conceptualizations in English and French
The French conceptualization is: surface
The French translation of the sentence follows:
 : Marc-Alain est sur le tableau

 CPU time : 101.34     GC time : 0.00
```

6.3.6.5 Reference object is a part of space or environment and can hide

In this case English and French use a different conceptualization. Compare with sections 6.3.6.3 and 6.3.6.4.

```
 : The boys are in the fog

***** INITIALIZATION *****

The located and the reference objects are not lexically ambiguous.

In English, 'fog', when used as a reference object, can be conceptualized as:
 C1.  SURFACE if its highlighted property is one of the following:
    1. part-of-object-support
```

```
 C2. an EMPTY VOLUME if its highlighted property is one of the following:
    1. environment-in-which-one-can-hide-cannot-hide-or-which-has-a-tiny-width
       part-of-space-or-environment


**** DERIVATION OF THE OBJECTIVE MEANING ****

Given the ideal meanings of 'in', only set C2 is selected.

For scene SC2.1 to be built, the following conditions must be satisfied.
 A. The situation must be normal
       * The size check is satisfied.
       * The materiality check is satisfied.
 --- No condition involving the located object needs to be checked ---
Because all the conditions are verified, scene SC2.1 is built.

***** TRANSLATION INTO FRENCH *****

The description of scene SC2.1 follows:
'Object in environment-in-which-one-can-hide-cannot-hide-or-which-has-a-tiny-width'
Both English and French use conceptual representation C2.1 to express this scene.
The French translation of the sentence follows:
 : Les garcons sont dans le brouillard

 CPU time : 77.99     GC time : 0.00
```

6.3.6.6 Reference object is the air seen as a location

In this case English and French use a different conceptualization.

```
 : The balloon is in the air

***** INITIALIZATION *****

The located and the reference objects are not lexically ambiguous.

In English, 'air', when used as a reference object, can be conceptualized as:
 C1.  SURFACE if its highlighted property is one of the following:
    1. part-of-object-support
 C2. an EMPTY VOLUME if its highlighted property is one of the following:
    1. air-seen-as-a-location


**** DERIVATION OF THE OBJECTIVE MEANING ****

Given the ideal meanings of 'in', only set C2 is selected.

For scene SC2.1 to be built, the following conditions must be satisfied.
 A. The located object must be light or must be able to fly
       * This condition is verified because a balloon is light or can fly.
 B. The situation must be normal
       * The size check is satisfied.
       * The materiality check is satisfied.
Because all the conditions are verified, scene SC2.1 is built.
```

```
***** TRANSLATION INTO FRENCH *****

The description of scene SC2.1 follows:
'Object in air-seen-as-a-location'
The scene is expressed with different conceptualizations in English and French
The French conceptualization is: functional-container
The French translation of the sentence follows:
 : Le ballon est en l' air

 CPU time : 109.72     GC time : 0.00
```

6.4 The erroneous and abnormal input sentences

Now that we have presented runs of valid sentences, we will illustrate the performance of the system on invalid or abnormal sentences. As described in chapter 5, we distinguish between several cases of errors: lexical errors, where a word input is not listed in the lexicon; errors of conceptualization where the reference object cannot be conceptualized in any of the ways it is required to by the preposition; and usage errors or abnormalities, where the conditions involving the located object are not satisfied (usage error) or when the normality conditions are not satisfied (abnormality). The usage errors and abnormalities are handled together. See the discussion of chapter 5 about whether or not these two types of error or abnormality should be distinguished.

This section illustrates the four different cases of errors or abnormality with examples.

6.4.1 Lexical error

We show one example of lexical error. In this example. 'bboy' is not a word of the lexicon.

```
 : The bboy is in Paris

Run aborted !
Lexical error in the source sentence !

 CPU time : 13.13     GC time : 0.00
```

6.4.2 Conceptualization error

We show one example of an error of conceptualization. The use of 'at' requires 'shade' to be conceptualized as a POINT. However, 'shade' cannot be conceptualized as a POINT in English (note that it could be in French).

```
 : Joshua is at the shade

***** INITIALIZATION *****

The located and the reference objects are not lexically ambiguous.

In English, 'shade', when used as a reference object, can be conceptualized as:
 C1.  SURFACE if its highlighted property is one of the following:
    1. part-of-object-support
 C2. an EMPTY VOLUME if its highlighted property is one of the following:
    1. part-of-space-or-environment
       environment-in-which-one-can-hide-cannot-hide-or-which-has-a-tiny-width
Run aborted !
Error of conceptualization of the reference object in the source sentence !

 CPU time : 34.68     GC time : 0.00
```

6.4.3 Usage or abnormality errors

We show, in this order, two cases of usage error and one case of normality error (in which the materiality check is not satisfied).

In the first example, two scenes are attempted. However, in both cases, the conditions involving the located objects are not verified: in SC1.1, Canada is not part of Maxime and in SC1.2, Maxime is not atmospheric. No scene is therefore built.

```
 : Maxime is on Canada

***** INITIALIZATION *****

The located and the reference objects are not lexically ambiguous.

In English, 'Canada', when used as a reference object, can be conceptualized as:
 C1.  SURFACE if its highlighted property is one of the following:
    1. part-of-object-support
    2. uncovered-geographic-entity
 C2. a BOUNDED-SURFACE if its highlighted property is one of the following:
    1. bounded-surface


**** DERIVATION OF THE OBJECTIVE MEANING ****

Given the ideal meanings of 'on', only set C1 is selected.

For scene SC1.1 to be built, the following conditions must be satisfied.
 A. The located object must be a part of the reference object
       o This condition is not verified because a Canada is not part of a Maxime
 B. The situation must be normal
       * The size of the objects does not matter.
       * The materiality check is satisfied.
Scene SC1.1 cannot be built because all the conditions are not satisfied.
Conceptual representation C1.1 cannot be instantiated.

For scene SC1.2 to be built, the following conditions must be satisfied.
 A. The located object must be atmospheric
       o This condition is not verified because a Maxime is not atmospheric.
 B. The situation must be normal
       * The size check is satisfied.
       * The materiality check is satisfied.
Scene SC1.2 cannot be built because all the conditions are not satisfied.
Conceptual representation C1.2 cannot be instantiated.

***** TRANSLATION INTO FRENCH *****
Run aborted
Usage error or abnormality in the source sentence !

 CPU time : 147.95     GC time : 0.00
```

In the second example, again two scenes are attempted. However, the conditions involving the located object of each scene are again not satisfied: in SC2.1, the air is not a constituent of a glass, and in SC3.1, a glass is not hermetic to the air.

```
 : The air is in the glass

***** INITIALIZATION *****

The located and the reference objects are not lexically ambiguous.

In English, 'glass', when used as a reference object, can be conceptualized as:
 C1.  SURFACE if its highlighted property is one of the following:
    1. part-of-object-support
 C2. a FULL VOLUME if its highlighted property is one of the following:
    1. object-apprehended-as-a-whole
 C3. an EMPTY VOLUME if its highlighted property is one of the following:
    1. well-defined-low-container


**** DERIVATION OF THE OBJECTIVE MEANING ****

Given the ideal meanings of 'in', only sets C2 and C3 are selected.

For scene SC2.1 to be built, the following conditions must be satisfied.
 A. The located object can be a constituent of the reference object
       o This condition is not verified because a air is not a constituent
         of a glass
 B. The situation must be normal
       * The size check is satisfied.
       o The materiality check is not satisfied.
Scene SC2.1 cannot be built because all the conditions are not satisfied.
Conceptual representation C2.1 cannot be instantiated.

For scene SC3.1 to be built, the following conditions must be satisfied.
 A. The reference object must be hermetic to the located object
       o This condition is not verified.
 B. The situation must be normal
       * The size check is satisfied.
       o The materiality check is not satisfied.
Scene SC3.1 cannot be built because all the conditions are not satisfied.
Conceptual representation C3.1 cannot be instantiated.

***** TRANSLATION INTO FRENCH *****
Run aborted
Usage error or abnormality in the source sentence !

 CPU time : 93.14     GC time : 0.00
```

In the third example, it is the materiality check of the unique scene attempted that is not satisfied: shade is not material.

```
 : The shade is in the sun

***** INITIALIZATION *****

The located and the reference objects are not lexically ambiguous.

In English, 'sun', when used as a reference object, can be conceptualized as:
 C1.  SURFACE if its highlighted property is one of the following:
    1. part-of-object-support
 C2. an EMPTY VOLUME if its highlighted property is one of the following:
    1. part-of-space-or-environment


**** DERIVATION OF THE OBJECTIVE MEANING ****

Given the ideal meanings of 'in', only set C2 is selected.

For scene SC2.1 to be built, the following conditions must be satisfied.
 A. The situation must be normal
       * The size check is satisfied.
       o The materiality check is not satisfied.
 --- No condition involving the located object needs to be checked ---
Scene SC2.1 cannot be built because all the conditions are not satisfied.
Conceptual representation C2.1 cannot be instantiated.

***** TRANSLATION INTO FRENCH *****
Run aborted
Usage error or abnormality in the source sentence !

 CPU time : 59.74     GC time : 0.00
```

6.5 Summary

We will now evaluate the results of the system and discuss the cases that it cannot handle.

6.5.1 Evaluation of the system

We first recall the method used by the system. The system uses the English conceptualization of the input sentence to derive its objective meaning, and then finds the French conceptualization associated with this objective meaning for finally translating the input sentence into French. The process of finding the objective meaning of the sentence shows some level of 'understanding' in the system.

As demonstrated by the runs presented in this chapter, the system can handle the translation of different types of sentences. Some of the sentences illustrated in section 6.3 assume the same conceptualization in English and French, others, a different one. The similarity or dissimilarity of the conceptualization is recognized by the system and indicated in the output.

In order to derive the objective meaning of a sentence, the system uses different types of world knowledge, as demonstrated by conditions A and B of the scene building attempts of each run. This knowledge involves social functional knowledge about objects and natural properties about objects, such as their sizes, shapes, constituency, capacity to adhere, capacity to hide, etc.

In addition, as shown in sections 6.3 and 6.4, the system is able to detect all the types of ambiguity and errors described in chapter 5. Once these ambiguities and errors are detected, the system is able to output the appropriate messages in each cases, and handle the problem as discussed in chapter 5.

6.5.2 Cases the system cannot handle

Section 6.3 and 6.4 deal with examples of sentences that are satisfactorily processed by the system (even though section 6.4 deals with erroneous sentences). However, not all sentences can be processed satisfactorily. This is the case, for example with the four following sentences: "The boy is at school", "The boy is at the school", "The boy is in school", and "The boy is in the school", where the meaning of the sentence is affected in part by the presence or the absence of the article. We did not implement a system of recognition that interprets the presence or the absence of the article in all cases.

Other sentences that failed are, for example, "The girl in the shirt" and "The tree in the road". In these cases the failure is due to an error in the program.

Another type of problem is caused by the fact that our size scale is not precise enough: some sentences that should not be accepted, are, because of this problem. For example, the sentence "The plane is on the bus" is accepted by our system, whereas it should not

be. This could be corrected with a more precise scale.

Even though some sentences are not successfully processed by the system, we believe that it is not due to the approach we followed: we believe that these unsuccessful runs are caused by implementation errors in the program or by inadequate or lack of knowledge in the knowledge base. We think that all these problems could be overcome without the approach being modified.

Chapter 7

Conclusion

Now that our approach has been presented in its entirety, we will conclude this thesis by discussing some areas for future work and then giving a summary of our research.

7.1 Future work

In this section, we discuss two types of future work: work whose focus is to improve our current approach while not modifying its main goals, and work whose focus is to improve our current approach by adding to it new but related goals. In both cases, the suggestions for future goals are given both at the theoretical and the implementation levels.

7.1.1 Future work that does not modify current goals

In the framework of extensions of our work that do not modify our set of current goals, we discuss six areas in which such extensions are possible.

In the first area, the extension consists of augmenting the knowledge base. Our work has, as one of its main goals, the systematization of Grimaud's (1988) theory. We have already checked Grimaud's theory against many different types of cases (as presented in chapters 4 and 6) and decided that the approach seemed satisfying. However, our corpus remains limited. Increasing it would allow us to draw further and more definitive conclusions.

The second area for future work concerns the organization of the knowledge needed by the system. As discussed in chapter 3, some of this knowledge is stored in the knowledge base and some of it is stored in procedural form. The decision of how the knowledge should be organized was made on the basis of the implementation of this particular system: the information that needs to be accessed several times during a run is stored in the knowledge base, whereas the knowledge that needs to be accessed only once during a run is stored procedurally. The reorganization of this knowledge taking into account, on the one hand, what part of it is fundamental to the process of translation and what part of it is not, and, on the other hand, what part of it might be needed for other tasks in artificial

intelligence and what part of it might not, would lead to a more meaningful organization.

The third area of improvement is algorithmic. Currently, the algorithm builds a lot of unnecessary conceptual information. The algorithm would be more efficient if we restricted the building of the conceptual information to a minimum. This type of improvement should be made at the COB level. Using information about the preposition used in the sentence prior to building conceptual representations is one important action that can be taken towards this improvement.

In the fourth area for future work the extension consists of investigating the question about distinguishing between abnormalities and usage errors, which was discussed in chapter 5. As mentioned in chapter 5, even though we distinguish between abnormalities and usage errors, we are not sure that such a distinction should be made. In fact, we tend to think that there are different levels of abnormalities — some sentences are more abnormal than others — but no usage errors. Investigating this question would allow us to extend our system to contexts other than the normal one, by having a system able to deduce whether certain sentences are normal (what the system currently does), and if not, in what contexts they could be valid.

The last two areas for future work are external to the field of artificial intelligence, and belong to the domains of linguistics and psychology. These two areas of improvements are linked to each other, the first one being a prerequisite for the second. Improvements in the first of these areas consists of distinguishing between cases of the uses of topological prepositions that are derived by cognitive rules and cases that are simply conventional. We believe that certain uses are simply conventional, such as the case that involves proper names of countries and continents in French (see the discussions in chapters 4 and 6), whereas other cases are cognitively rooted (even though some level of convention is always present). Such an improvement would be the basis for our last area for future work which consists of concentrating on the cognitive cases and attempting to find fewer and more general cognitive rules than the current ones, able to account for the different uses of topological prepositions.

7.1.2 Future work that modifies current goals

In the framework of extensions of our work that add new but related goals to the current study, we discuss four areas in which such extensions are possible.

The first area concerns the corpus. We have presented a systematization of the approach of Grimaud on the three basic topological prepositions. We would like to extend this systematization to other types of topological prepositions, such as the projective ones, for example. As well, we would like to be able to process sentences with verbs other than 'to be', and more complicated syntax structures, since we think that these types of modifications might have an effect on the uses of topological prepositions.

The second area of improvement consists of developing better representations for the

objective meanings of the sentence. This was not needed by our system, which limits itself to performing translation, because only the conceptual representations of the reference objects turned out to be relevant. However, in order to extend our system to perform other tasks, such an improvement would be in order. One of the tasks we have in mind, for which such an improvement would be necessary, is the linking of visual information with linguistic information.

In the third area, the improvement consists of adding a pragmatic component to our system that would be able to derive more precise meanings of the sentences processed by understanding the possible goals of the utterances according to the topological preposition used.

The last area for future work is not directly related to artificial intelligence, or even cognitive sciences. It consists of explaining the conventions in language, which, in some cases, probably come from historical uses and, in others, are simply arbitrary. We think that this area is related to our work because we believe that studying the evolution of languages would give us clues concerning language acquisition, which is an area investigated in both artificial intelligence and cognitive sciences.

7.2 Summary

Now that we have presented suggestions for future work, we will discuss the contributions of our current research. We first summarize our main contributions and results, and then give a detailed outline of our approach.

7.2.1 Main contributions and results

We have designed a system for the translation of topological prepositions from English into French. Even though English and French seem close in their uses of topological prepositions, many cases can be found where they differ. These cases are numerous and complex. Our system uses systematicities in languages that were studied by linguists and psychologists to solve this problem, and does so in a satisfying way. Our work, however, does not limit itself to the problem of translation: the system and the knowledge representation entities that we designed are also important for other tasks of computational linguistics in particular, and artificial intelligence in general.

In addition to translating sentences, our system is, first of all, able to derive conceptual information. Conceptual information is also important for cognitive tasks other than translation. Vision systems, for example might use conceptual information. We have designed an entity, the *conceptual representation of objects*, that represents the way objects can be conceptualized, given certain conditions. This type of entity was designed for our particular system, but nonetheless, could be used in, or at least serve as a model for, other types of cognitive tasks, as well.

Second, our system is able to derive objective meanings of sentences. Moreover, it is

able to distinguish between subtle differences of meanings in sentences, leading to the detection of ambiguity. This shows a certain level of "understanding". This level of "understanding" also allows our system to distinguish between erroneous, abnormal, and normal sentences. Detecting ambiguity and rejecting non-acceptable input sentences are important features of a computational linguistic system.

Finally, our work is relevant to cognitive sciences, because it is sensitive to studies of psychology attempting to determine the way people behave. We used results generated from the field of psychology as constraints on our system. As well, whenever faced with several possibilities, we chose psychology results over others.

Note that instead of our current system, we could have designed a system that would have simply performed "blind" translation without attempting to "understand" the meaning of sentences. This hypothetical system might have worked in translating sentences from English to French satisfactorily (even though, there might be too many cases to take into consideration for such an approach to succeed). However, it would have been shortsighted from two points of view. First, in the domain of translation, such a system would be using the direct method of translation, which would make this system non-extendable to other languages, whereas our system, based on the transfer method of translation, could be valid for other languages, as well.[1] In addition, only minimal work would be needed to extend our current system to these other languages: the ideal meanings of the prepositions in the DEC and the conceptual rules of the COB are the only parts of the program that would require modifications (together with the lexicon, the parser, and the generator, of course). Second, such a system as the hypothetical one would not have been interesting from an artificial intelligence point of view since it would have been restricted to a unique task — i.e., that of translation, whereas our approach has led us to design entities that might be useful for other types of important tasks, not specific to machine translation.

7.2.2 Detailed summary of our work

7.2.2.1 Our work in relation with previous research

Our work takes as a starting point the theory of Herskovits, whose main goal is an attempt to account for the lexical meanings of spatial prepositions. Herskovits's theory has uncovered certain rules that account for the lexical meaning of spatial prepositions. A large number of uses, however, remained unexplained. Herskovits labels these uses conventional. Our principal goal was to try and go one step further in order to find some orderliness in these assumed 'conventional uses'. The way we chose to pursue this goal was to study the different uses of topological prepositions in a multi-lingual frame, which, we thought, would give us more insight into this phenomenon. Given the scope of our study, we had to reduce the set of prepositions studied to a small one: the set

[1] Note that we are talking about languages, such as European languages, presenting uses of spatial prepositions similar to English and French. We are not claiming anything about other types of languages which we did not study.

of basic topological prepositions. As well, we studied the phenomenon within only two languages: English and French. These languages have the advantage of being closely related to each other to the extent that a meaningful comparison can be made; yet they are different enough to provide an interesting basis for a comparison.

In order to pursue our goal of understanding the uses of topological prepositions, we have adopted the theory of Grimaud, which attempts to explain how these assumed 'conventional uses' are organized. Grimaud's study is a comparative study of English and French in their uses of spatial expressions, relying heavily on Lakoff & Johnson's (1980) theory of metaphors. Grimaud's theory assumes that real world objects are conceptualized in different ways in English and French and that these differences explain the differences in their uses of spatial prepositions. We have adopted this theory and applied it systematically to sentences or expressions of English and French, in order to show its plausibility. We did not get rid of the notion of conventionality in languages, but were able to shift it from the lexical meaning of prepositions to the lexical meaning of nouns. We believe that this shift is very important because it allows us to deal with the use of topological prepositions as a general task if we adopt Lakoff & Johnson's (1980) theory — that the human conceptual system is metaphorical in nature. If we adopt this theory, then all cognitive tasks must use metaphors (or conceptualizations). Determining the uses of topological prepositions constitutes only one of these tasks. With Herskovits's theory, the uses of topological prepositions is conventional. It must therefore be treated as a special task of cognition which is not at all related to others. Our approach seems more plausible. In addition, some linguistic questions seem easier to answer given our approach than given Herskovits's. With Herskovits's scheme — in which a large number of the uses of spatial prepositions remain conventional — there is no way to account for the origin of this convention: it seems that it is totally arbitrary, as if some languages randomly adopted one convention, whereas others adopted other ones. By shifting the convention from the lexical meaning of prepositions to the conceptualizations of objects, we think that we are able to provide some linguistic explanations for these uses. Even though we did not research this area — which is more the domain of a linguist than a computer scientist — we believe that the conceptualization of objects in a culture originates with the use of this object, its history, its status, its shape, etc., in that culture. When considering all these factors, the different uses of topological prepositions seem systematic and meaningful. As well, considering these factors allows us to account for the evolution of the uses of prepositions within a language, since the shape or uses of objects evolves with time, and, as a result, so do their conceptualizations.

We have built a computational model for Grimaud's theory which we have tested on many examples, and in particular the ones listed by Herskovits as due to the conventional uses of the basic topological prepositions.

7.2.2.2 Detailed summary of our work

We will now present the main results and contributions of our work. We have implemented Grimaud's (1988) theory, using many of the ideas of Herskovits's, in order to

build a system which performs the translation from English to French of sentences and expressions involving one of the basic topological prepositions. We will first review the main systems of representation that we have designed and used, and then will review the main processes of the system.

The most central entity of Herskovits's that we used is the *ideal meaning*. Like Herskovits, we assumed that each preposition has an ideal meaning. In addition, we assumed that French prepositions have the same ideal meanings as their English equivalent. We, however, used the notion of ideal meanings differently from Herskovits. From the ideal meaning of the prepositions, Herskovits derives new uses of these prepositions through the mappings called *deviations* and *shift*. Our technique differs in that in our model, the prepositions do not depart from their ideal meaning; instead, the objects are conceptualized in different ways according to the situation, causing the prepositions to be used in different ways, as well. Another difference between the two systems is that in our system each preposition may have more than one ideal meaning, whereas in Herskovits's, each preposition has only one ideal meaning.

In addition to the aforementioned system of representation for the prepositions, we developed a system for the representation of objects. This system is composed of two main parts: the first taking care of the *objective representation* of objects, the second taking care of the *conceptual representation* of objects. To each objective representation of objects there corresponds at least one conceptual representation.

The objective representation of objects includes information relevant to their spatial relations, such as their shapes, functions, sizes, etc. The objective representation was inspired by the work of Peters et al. (1987ab; 1988).

The other representations, the conceptual representations, are composed of two parts: a metaphor (or conceptualization) in which an object can be conceptualized; and a set of conditions in which such a metaphor can be instantiated (this information is available from the COB). The conceptual representations of objects was designed by us. The conceptual representations of objects are built during the run of the system, using objective information stored in the knowledge base and a set of conceptual rules, (procedurally) stored in the COB. This set of conceptual rules is fundamental to the system. Given objective and syntactic information about an object, these rules derive the conceptual representations of that object.

We will now review the main processes of the system.

The first important procedure to be called during a run is the COB. The COB uses objective information about the located and reference objects (which were identified by the SLU in an earlier step) together with conceptual rules, and builds all the conceptual representations of objects. Some of these conceptual representations are then eliminated by the DEC, which uses the ideal meanings of the preposition for carrying out this task. The DEC proceeds directly after the COB has built all the conceptual representations of

the two objects of the sentence.

The next important procedure is the SB. The SB has two roles. First, it eliminates further some of the conceptual representations built by the COB which remain after the elimination carried out by the DEC, and second, it builds the different objective meanings of the sentences, when the sentences are not erroneous or abnormal. The first task is carried out both by checking whether the located and reference objects are compatible with each other in the possible scenes represented by the sentence and by checking whether the sentence is normal. These two operations use, for the most part, world knowledge about objects. The second task consists of gathering the relevant information into an entity — the objective meaning of the sentence.

The next important procedure is the MATCHER. This procedure and those subsequent to it might be called several times during a run, in the case of conceptual ambiguity. The role of the MATCHER is to find the conceptualizations that are used in French to express the objective meaning currently being processed. Once the MATCHER finds a French conceptualization, the French preposition corresponding to this conceptualization is found by the FPD, using a simple look-up table, and the French sentence is generated by the TLG.

The entire procedure just described is then repeated (except for the SLU procedure) if the sentence has a complex lexical ambiguity at the level of the located or the reference object.

In addition to translating sentences, the system detects ambiguity and errors or abnormalities.

The performance of the system is good with regard to the range of cases it can treat and its reliability. The following is a list of all the types of sentences it can translate: sentences which use the same conceptualization in English and French and sentences which use different ones; sentences with simple or complex lexical ambiguity; sentences with simple or complex conceptual ambiguity (sentences with several ambiguities of the same or different types, at once are also translated in a satisfying way). As well, the system can detect sentences with the following type of errors in them: lexical errors, errors of conceptualization, usage errors[2], and abnormal sentences.

There is one problem with our system, however: its performance is poor in computing time. This problem is independent of us: it comes from the fact that a lot of world-knowledge information is needed for the task of "understanding" and translating sentences. This world knowledge needs to be stored in the semantic network prior to the run of the system and the storing of a large number of SNePS nodes in the semantic network slows the system down tremendously.

[2]See section 7.1.1 for a discussion of whether or not this type of error should be considered an error.

We will now conclude this thesis with a brief evaluation of our work. We have accomplished three tasks. First, we have systematized Grimaud's (1988) theory by building a computational system based on his theory, able to process different types of sentences. By accomplishing this, we have reinforced the theory.[3] Second, we designed an important entity — the conceptual representation of objects. This entity is important in the field of artificial intelligence, since other cognitive tasks might need conceptual knowledge and could use entities similar to the one we designed. Finally we met our primary goal, which was to design a system of translation which works satisfactorily in the framework we defined.

[3] However, further research in psychology and linguistics should be undertaken, for better conclusions.

Bibliography

[Bennett 1968] Bennett, D. C. 1968. *English Prepositions: A Stratificational Semantics.* London: Longman.

[Bennett 1975] Bennett, D. C. 1975. *Spatial and Temporal Uses of English Prepositions: An Essay in Stratificational Semantics.* London: Longman.

[Bierwisch 1967] Bierwisch, M. 1967. "Some Semantic Universals of German Adjectivals". *Foundation of Language,* vol. 3.

[Chomsky 1982] Chomsky, N. A. 1982. *Lectures on Government and Binding: The Pisa Lectures (Studies in Generative Grammar),* second edition. Dordretch: Foris Publications.

[Clark 1973] Clark, H. H. 1973. "Space, Time, Semantics, and the Child". *Cognitive Development and the Acquisition of Language,* ed. T.E. Moore, 65-110. New York: Academic Press.

[Cooper 1968] Cooper, G. S. 1968. *A Semantic Analysis of English Locative Prepositions,* Bolt Beranek and Newman report No. 1587. Springfield, Va.: Clearinghouse for Federal Scientific and Technical Information.

[Dahlgren 1988] Dahlgren, K. 1988. *Naive Semantics for Natural Language Understanding.* Kluwer Academic Publishers.

[Gougenheim 1938] Gougenheim, G. 1938. *Système grammatical de la langue française.* Paris: D'Artray.

[Grimaud 1988] Grimaud M. 1988. "Toponyms, Prepositions, and Cognitive Maps in English and French". *Journal of the American Society of Geolinguistics.* vol. 14, pp. 54-76.

[Hawkins 1983] Hawkins, B. 1983. *Semantics of English Spatial Prepositions.* Ph.D. Thesis. San Diego, California: University of California. Thesis pre-published by L.A.U.D.T., Linguistic Agency of the University of Duisburg, previously Trier.

[Herskovits 1986] Herskovits, A. 1986. *Language and Spatial Cognition: An Interdisciplinary Study of the Prepositions in English.* Cambridge: Cambridge University Press.

[Herskovits 1989] Herskovits, A. 1989. *The Linguistic Expression of Spatial Knowledge.* Duisburg: L.A.U.D. 1989. Report no. A 248.

[Kuipers 1978] Kuipers, B. 1978. "Modeling Spatial Knowledge". *Cognitive Science,* vol. 2, 129-53.

[Lakoff & Johnson 1980] Lakoff, G. & Johnson, M. 1980. *Metaphors we Live by.* Chicago: University of Chicago Press.

[Lakoff 1982] Lakoff, G. 1982. "Categories and Cognitive Models". *Cognitive Science Program,* Berkeley, California: University of California.

[Lakoff 1987] Lakoff, G. 1987. *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind.* Chicago: University of Chicago Press.

[Leech 1969] Leech, G. N. 1969. *Towards a Semantic Description of English.* Bloomington: Indiana University Press.

[Mark, Svorou, & Zubin 1987] Mark, M., Svorou, S. & Zubin, D. 1987. "Spatial Terms and Spatial Concepts: Geographic, Cognitive, and Linguistic Perspectives". *Proceedings of the International Geographic Information Systems (IGIS) Symposium: The Research Agenda,* vol. 2, pp. 101-112.

[Mark 1990] Mark, D. 1990. ""The House on the Lake": Towards a Cross-Linguistic Computational Model of a Geographical Relation". Unpublished manuscript, Research Initiative 2, National Center for Geographic Information and Analysis, State University of New York, Buffalo, NY.

[Mervis & Rosch 1981] Mervis, C.B., Rosch, E. 1981. "Categorization of Natural Objects", *Ann. Rev. Psychol.,* vol. 32, pp. 89-115.

[Miller & Johnson-Laird 1976] Miller, G. A. & Johnson-Laird, P. N. 1976. *Language and Perception.* Cambridge, Mass.: Harvard University Press.

[Peters & Shapiro 1987a] Peters, S. & Shapiro, S. C. 1987a. "A Representation for Natural Category Systems". *Proceedings of the Ninth Annual Conference of the Cognitive Science Society,* pp. 379-390.

[Peters & Shapiro 1987b] Peters, S. & Shapiro, S. C. 1987b. "A Representation for Natural Category Systems". *Proceedings of the Tenth International Joint Conference on Artificial Intelligence,* pp. 140-146.

[Peters, Shapiro & Rappaport 1988] Peters, S., Shapiro, S. C. & Rapaport W. J. 1988. "Flexible Natural Language Processing and Roschian Category Theory". *Program of the Tenth Annual Conference of the Cognitive Science Society,* pp. 125-131.

[Retz-Schmidt 1988] Retz-Schmidt, G. 1988. "Various Views on Spatial Prepositions". *A.I. Magazine* vol. 9, no. 2, pp. 95-105.

[Rosch et al. 1976] Rosch, E., Mervis, C.B., Gray W.D., Johnson, D.M., Boyes-Braem, P. 1976. "Basic Objects in Natural Categories". *Cognitive Psychology,* vol. 8, pp. 382-439.

[Rosch 1977] Rosch, E. 1977. "Human Categorization". *Advances in cross-cultural psychology,* vol. 1, ed. N. Warren, 1-49. London: Academic Press.

[Rosch et al. 1978] Rosch, E., Lloyd, B.B. 1978. *Cognition and Categorization.* Hillsdale, N.J.: Lawrence Erlbaum Associates).

[Shapiro 1979] Shapiro S. C. 1979. "The SNePS Semantic Network Processing System". *Associative Networks,* ed. N.V. Findler, pp. 179-203. New York: Academic.

[Shapiro 1982] Shapiro, S. C. 1982. "Generalized Augmented Transition Network Grammars for Generation from Semantic Networks". *American Journal of Computational Linguistics,* vol. 8, pp. 12-25.

[Shapiro & Rapaport 1987] Shapiro, S. C. & Rapaport, W. J. 1987. "SNePS Considered as a Fully Intensional Propositional Semantic Network". *The Knowledge Frontier,* ed. N. Cercone & G. McCalla, pp. 262-315. New York: Springer-Verlag.

[Shapiro 1989] Shapiro, S. C. 1989. *SNePS-2.1 User's Manual.* State University of New York at Buffalo.

[Slocum 1988] Slocum, J. 1988. "A Survey of Machine Translation: its History, Current Status, and Future Prospects". *Machine Translation Systems,* ed. J. Slocum, pp. 1-47. Cambridge: Cambridge University Press.

[Smith & Medin 1981] Smith, E.E., Medin, D.L., 1981. *Categories and Concepts.* Cambridge, Mass.: Harvard University Press.

[Talmy 1983] Talmy, L. 1983. "How Language Structures Space". *Spatial Orientation* ed. H.L. Pick & L.P. Acredolo, chapter 11. New York: Plenum Press.

[Vandeloise 1986] Vandeloise, C. 1986. *L'espace en français. Sémantique des prépositions spatiales.* Paris: Seuil.

[Wade 1983] Wade, T. 1983. *Prepositions in Modern Russian.* Durham modern language series: University of Durham.