



**HAL**  
open science

# Recognition of Textual Inference a State of the Art

Djallel Bouneffouf

► **To cite this version:**

Djallel Bouneffouf. Recognition of Textual Inference a State of the Art. [Research Report] IRIT. 2013.  
hal-00820436

**HAL Id: hal-00820436**

**<https://hal.science/hal-00820436v1>**

Submitted on 7 Apr 2014

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# State of the Art on the Recognition of Textual Inference

Djallel Bouneffouf,

Department of Computer Science, Telecom SudParis, UMR CNRS Samovar, 91011 Evry  
Cedex, France

{Djallel.Bouneffouf} @ it-sudparis.eu

**Abstract.** This paper is devoted to the study of textual inference, the presentation of different applications of the Recognition of Textual Inference (RTE) and the main levels of textual inferences. We show the stages of development of the challenge Pascale RTE, which is implemented to assess the advances of research groups in this field, and we conclude with the presentation of some inference methods used by research groups evaluated in the challenge Pascal RTE.

**Keywords:** Natural language processing, inference engine, machine learning, artificial intelligence.

## 1 Introduction

Historically, the first major work in the field of NLP has focused on machine translation, with, in 1954, the development of the first automatic translator (very basic). Some Russian phrases selected in advance, were translated to English.

Since 1954, heavy funding has been invested and much research has been launched. The main work presented then related the manufacture and handling of electronic dictionaries as translation techniques are essentially translate word by word, then with a possible rearrangement of the order of words.

This simplistic view of translation led to the following famous example: the sentence *The spirit is willing the flesh is weak goal*

was translated into Russian and then translated back into English. This gave something like: *Vodka is strong but the meat is rotten*

What emerges from this example is that many contextual knowledge (ie on the situation described) and encyclopedic (ie on the world in general) are needed to find the correct translation of a word (eg by spirit which, depending on the context can be translated as spirit or as alcohol).

Posing as a conjecture that every aspect of human intelligence can be described in sufficient detail for a machine the simulates the most prominent figures of the time (John McCarthy, Marvin Minsky, Allan Newell, Herbert Simon) y discuss the possibilities of creating computer programs that behave intelligently, and in particular who are able to use language.

Today, the field of natural language processing is a very active field of research. Many industrial applications (machine translation, information retrieval, natural language interfaces), which are beginning to reach the general public, are there to show the importance of the progress made but also the progress that remains to be done.

## 2 Treatment levels

We introduce in this section the different levels of treatments required to achieve a complete understanding of a sentence in natural language. These levels correspond to modules that should be developed and to cooperate in the framework of full implementation of language processing.

We consider as an example the following statement:

(1) *The president of the temperance eating an apple with a knife*

We envision successive treatments to be applied to this statement to automatically reach its fullest understanding. We will successively:

- Identify the lexical components and their properties is the processing step **lexical** ;
- Identify components (group) of the highest level, and relationships (dominance) they maintain them: this is the stage of treatment **syntactic** ;
- Construct a representation of the meaning of this statement, by associating each concept discussed an object or action in a world reference (real or imagined) is the processing step **semantic**.
- Finally identify the function of the statement in the context of the particular situation in which it was produced: this is the stage of treatment **pragmatic**.

### 2.1 The lexical level

The purpose of this processing step is to move from atomic forms (tokens) identified by the word segmenter (Nugues, 2006), which mean to recognize in each string one (or more) unit (s) language (s), with characteristics (its meaning, its pronunciation, its syntactic properties, etc.).

In the example (1) The lexical identification step should lead to a similar result to that given below, in which we can see in particular the ambiguity of such a form that President: this string is two forms of the verb chair (code and subjunctive), and a nominal form, and its pronunciation differs depending on whether it is a noun or a verb.

It is easily understood that for the most frequent words, like "the", the simplest solution is to look in the form (a glossary)<sup>1</sup> precompiled. In fact, that is what is happening, including the rarest forms, insofar as the use of compact representation formalisms allowing more efficient access (eg in the form of finite state automata), and increasing the size of memory makes it possible to manipulate large lexicons (of the order of hundreds of thousands of forms).

---

<sup>1</sup> In languageThe lexicon of a language all its forms lemmas or, more common, but less accurately, "all its words". Also in common usage, the term vocabulary is used more easily.

However, this solution does not solve all problems. Language is creation, and new forms arise every day, whether through borrowing from other languages (there has only been heard about teachers in other modules of the dominant computer!), Or more frequently by the application of regular processes creations of words that allow us to compose almost at will new forms immediately understood by all speakers of our language if I like reading Proust, can we not say that I emproustise me, as I become Proustian proustiste or proustophile and that tired, I dis-semproustise ... This phenomenon is not marginal, as it is recognized that even if one has a complete lexicon of French, about 5 to 10% of the words in a newspaper article at random will not appear in this lexicon. Purely lexical solution reaches its limits here, so you have to implement other approaches, so as to treat the forms off-lexicons.

## 2.2 The syntactic

The syntax is the study of constraints on the legal succession of forms that must be taken into account when trying to describe the sequences forming grammatically correct sentences: all strings of words do not form acceptable sentences (Ligauzat, 1994). Constraints describing the characteristics of a given language is through a grammar. Models and formalisms proposed in the context of natural language processing are particularly numerous and varied.

The syntactic level is the conceptual level involved in the calculation of the validity of certain sequences of words, grammatical or well-formed sequences. The importance of such treatment in a generation application is conceived, why it is essential that the machine generates correct statements. In an application of understanding, machine analysis of texts which are submitted to it, which we can assume that they are grammatical. Why, in this case, implement syntactic knowledge?

A primary motivation comes from the fact that the texts are not always grammatical, for example because of spelling mistakes. Parsing can therefore afford to choose between several corrections to incorrect sentence, but also be very useful to improve the outputs of an optical character recognition or even a system of speech recognition.

A second reason is that the entrance to the syntactic module is a series of labeled morphosyntactic forms, a form can have several different labels. A primary function of the syntactic module therefore is to disambiguate the following labels, eliminating sequences that correspond to statements grammatically invalid.

## 2.3 The semantic level

Intuitively, the semantics is concerned with the meaning of utterances (Yvon, 2007). A sentence like *The garden door eat the sky* Although perfectly grammatically correct, does not make sense in most contexts. But what is the meaning? For expression as *bottle right* in the sentence:

*Help yourself wine. No, not that one, take the bottle right.*

The meaning is the object (the concept) designated. In this example, the meaning depends heavily on the context: we need a representation of the scene to know what bottle is, so what brand it is.

For a predicate expression as *He ordered a 1982 Margaux* The direction can be represented by a predicate logic as  $\langle \text{demander}(\text{paul}, \text{chateau\_margaux\_82}) \rangle$ .

The again the context identification of such a predicate depends. The word order would in fact referred to another predicate if it were to command a vessel.

## 2.4 The pragmatic level

The pragmatic level is perfectly separable from the semantic level. While the semantics is concerned with the meaning of the statements, pragmatic concerns attitudes (truth, desirability, probability) that speakers adopt towards forward and logical operations that trigger these attitudes (Yvon, 2007).

Historically, some linguists have all called pragmatic language processing involving the context of utterance. This criterion has very little interest in that semantic processes are the same, the context intervene or not. However, there is a very important distinction, based on the notion of logical inference. Consider the following example:

(A) Peter, will you come to the dance tonight?

(B) Mary: I heard that Paul will be there!

The second sentence is interpreted as a negative response if it is known that *Marie Dislikes Paul*.

This interpretation is not semantic. From understanding the meaning of the intervention *Marie, Pierre* performs a logical inference using contextual knowledge, the enmity between Paul and Mary. Peter concludes that Mary does not want to go to the ball, ie it reconstructs Mary's attitude in relation to its own statement. This is not a conceptual construction, it is a logical operation. It therefore belongs to the pragmatic.

Techniques corresponding to this level of treatment are still poorly controlled. The pragmatic level, even if the techniques that correspond to it are not yet stabilized, appears less difficult to tackle the semantic level. Indeed, it seems that based on a fixed set of principles such as the principle of relevance, it is modeled correctly. Determining the argumentative intent of the author or speaker is essential in many applications, including the management of dialogue, text summarization, machine translation, the context-sensitive help systems or education, etc. . We therefore expect significant progress to this level in the coming years.

## 3 Difficulties of NLP: ambiguity

Natural language is ambiguous, and this ambiguity is reflected in the multitude of options for each relevant level of processing linguistic entities interpretations, as in the following examples:

### 3.1 Ambiguous graphemes (letters)

This ambiguity is in the process of orthographic encoding by comparing the pronunciation of *i* in *bed*, *pear* and *home*.

### 3.2 Ambiguity in the grammatical and semantic properties

Thereby *eat* is ambiguous both morpho-syntactically, since it corresponds to the indicative and subjunctive forms of the verb *eat*), but also semantically. Indeed, this form may as well refer (in a familiar style) to a set of conventional measures (such as sitting at a table, put a towel, use a variety of utensils, this possibly maintaining an interaction with another human ) with a final ingest food (which it does not require a direct object) vision and action of actually ingest a particular type of food (in which case it requires a direct object Direct ), etc.. Compare in effect:

- (a) Tomorrow, Paul eats with my sister.
- (B) Paul eats his bread with chocolate.

And the deductions that can be made from these two statements: from (a), one can reasonably conclude that Paul is sitting at a table, will have covered ... , While this is not necessarily true in the case of statement (b).

### 3.3 Ambiguity of the grammatical function of groups of words

The ambiguity is illustrated by the following sentence:

*he says the girl bike.*

In this example *Cycling* is either a complement to how to proceed (and it is *it* pedaling), a supplement maiden name (which is *it* which churns);

### 3.4 Ambiguity of the scope of quantifiers, conjunctions and prepositions

Thus, in *All my friends have had a drink*, we can assume that each had a different drink, but in *All witnesses heard a cry*, it is likely that it was the same cry for all witnesses. Similarly, when dogs and cats Paul is mentioned, the most natural interpretation is to understand Paul as a complement to the group name in cats and dogs that reading is much less natural in dogs Race and cats Paul;

### 3.5 Ambiguity in the interpretation of an utterance in context

We compare the "meaning" of *not*, in the following two exchanges:

- (A) *If I go to school tomorrow? Not (negation)*
- (B) *You're going to school tomorrow! No! (I do not believe).*

Indeed, the ambiguity is a major problem of NLP. To overcome the researchers created an area that centralize the problem and suggest methods of language processing at the lexical, syntactic and semantic level regardless of a given application aims. In what follows we will explore this area and these different applications.

## 4 Recognition of textual inference (TEN)

The ETR is a relatively new field of research in language processing (2005), which bring together the research in NLP to propose methods of language processing at the lexical, syntactic and semantic level independently of a particular application (summary aims to automatic system to answer questions or information retrieval).

The RTE is to automatically determine whether a text segment (H) is derived from another text segment (T) (Dagan et al, 05).

*Example:*

*T: "Amine 40 degrees **fever**His mother took him immediately to the**hospital** ".*

*H: "Amine **sick** ".*

In the example above, understand that the H segment is derived from the T segment is a simple deduction to humans, but the machine is quite another. To do this, researchers have proposed several approaches to solve the problem.

In the example, to say that H is inferred from the T system must link being sick (text H) with the word hospital and fever (T text) to deduce that there is inference.

In this section, we present the various applications of RTE, then we show the development stages of the challenge Pascale RTE was implemented to assess the advances of research groups in this field.

We develop in Section 2, the main levels of textual inferences and we conclude this chapter by presenting some inference methods used by research groups evaluated in the Easter challenge RTE.

### 4.1 Applications RTE

The inference between text segments is at the heart of many applications of automatic natural language processing (NLP). We describe in the following how the RTE helps in these areas:

#### **Information retrieval.**

The information retrieval is the science of searching for information in documents, databases, whether relational or networked via hyperlinks (Joachims, 2003).

The information retrieval is historically linked to information science and librarianship who have always had the desire to establish representations of documents in order to retrieve information through the construction area index . Computer technology has enabled the development of tools to process information and establish representation of documents at the time of indexing and retrieving information.

Approaches previously used were based on the search keywords in the text. The problem with these systems is that they do not take into account the relationships between key words or their meaning.

Example 1:

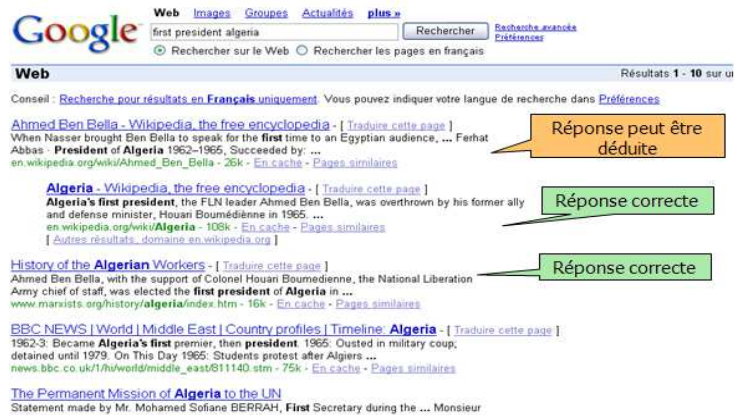


Figure 1.1: Example of search-based keyword

In this example (Figure 1.1) we see that a search engine powered base keyword as Google does what type of research and responds well to a simple question like "the first president Algeria" as a simple keyword search in the various documents can give a good answer to the user.

Example 2:



Figure 1.2: Example where search-based keyword does not work



In this example (Figure 1.2) we see that the use of key words alone can lead us to a document that has no relationship with our request and shows that the semantic inference is essential to search for information.

**The extraction of information.** Information extraction is to identify the precise a natural language text and to represent information in a structured format. For example, from a report on a car accident, an information retrieval system will be able to identify the date and place of the accident, the type of incident and the victims. This information can then be stored in a database to conduct further research or be used as a basis for the automatic generation of summaries (Kosseim., 2005).

The extraction of information is very useful in industry where operations extractions are performed daily by hand. We believe, for example, report processing spinning a monitoring agency, management dispatches a news agency, handling of insurance incident reports of a company, etc..

An information retrieval system automatically and quickly process large amounts of documents.

In this case the RTE gives its contribution in the detection of information.

#### **The question-answering system .**

Systems Questions / Answers are able to answer written in natural language searching the answer in a text corpus issues. They typically consist of a set of modules performing respectively an analysis of the issue, a search for relevant portions of documents and extracting the response using extractions patterns, or patterns in English (Nyberg et al 2002).

The system must identify the text segment which contains the response. The inference from the text T and H segment can help detect the segment that contains the answer.

Example:

H: "who is Ariel Sharon? "

T: "Israel's Prime Minister Ariel Sharon visited Prague."

The system first performs a transformation in the affirmative to the question "Ariel Sharon is Isreal's Prime Minister" and a comparison of the text segment T and H segment

If H is inferred from T as in the example then T is accepted as a segment containing the answer H.

#### **Automatic translation .**

Automatic translation means, strictly speaking, being fully translate a text with one or more computer programs, without a human translator has to intervene (Laurian and Mary, 1996). Machine translation is still very imperfect, and the generation of translation of a comparable quality of human translators is still something of a utopia.

To evaluate the performance of the machine, the RTE can compare the translation made by the machine that made by humans.

### **5.2.5) The automatic summarization**

Automatic summarization aims to make information retrieval considered an important text entered to construct, from this information, a new text output condensed. This new text avoids reading the entire source document.

The RTE is used to find information redundancies.

If a text segment infer another, one of the two will be deleted.

In particular it is interesting in applications that make the summary of several documents. If there are several documents that tell the same fact, one should be taken.

### **5.2.6) Acquisition of Paraphrases (AP)**

A paraphrase is the Saying with other words, in other words what is said in a text, a paragraph.

In this case the RTE is used to detect the inference between text and paraphrased the original text. As in the following example where the two sentences have the same meaning with just another arrangement of words in the sentence.

Example:

T: "This drug is sold only in Canada."

H: "The marketing of this drug is made in Canada only."

### **The challenge "of PASCAL Recognizing Textual Entailment"**

Pascal recognition of Textual Entailment is a competition that began in 2005. It takes place every year and its goal is to provide the community of NLP a new landmark to check progress in the recognition of textual inference, and compare the achievements of different research groups working in this field ( <http://www.pascal-network.org/Challenges/RTE/> ).

Following the success of the first new TEN TEN was held with 23 groups from around the world (compared to 17 for the first challenge) who presented the results of their systems. The representatives of the participating groups presented their work **PASCAL Challenges Workshop** in April 2006 in Venice, Italy.

The event was a success and the number of participants and their contributions to the discussion showed that the Textual Entailment is a rapidly expanding field. Already, the workshops led to an impressive number of publications in major conferences, in addition to ongoing work.

The steps taken to achieve the competition are:

Preparation of the corpus.

Establishment of measures evaluated.

In the following the mentioned steps are detailed.

### **5.3.1) The preparation of the corpus**

The first step to take is to create a corpus of text-hypothesis (TH) even small text segment, which corresponds to the information collected through the web in different fields.

The samples were collected manually for inference by human annotators.

The samples were divided into two types of corpus (**Corpus Development** and **Corpus**).

**The development corpus** is used at the beginning of a challenge to give users the opportunity to test their systems and make minor adjustments to prepare for the test.

**The corpus** is used for the final evaluation.

For RTE 1 The corpus consisted of 567 pairs (HT) for the development and 800 pairs for testing.

The choice of a large corpus is justified by the need for statistically significant results.

The corpus is collected in compliance with the various applications of natural language processing (QR, IR, IE, ... PP.) And collection of examples is given by level of inference:

The lexical, syntactic, logical analysis and knowledge of the world, and different levels of difficulty.

```
<pair id="754" value="True" task="CD">
  <t>
    Mexico City HAS very bad pollution problem Because the mountains
    around the city act as walls and block in dust and smog.
  </ T>
  <h> Poor air flow out of the mountain-walled Mexico City ag-
  gravates pollution.</ H>
</ Par>
Id: represents the number together.
Value: is the decision of the annotator (true or false).
Task: represents the type of the application is or inference.
```

**Figure 1.3: Example of annotated corpus**

The body should include 50% of a sample corresponding to TH real inferences and 50% false inferences. To do this, each example (TH) is judged true or false by the annotator creates example.

Then the example is evaluated by a second judge who evaluates the pairs of texts and assumptions, without being aware of their contexts.

Annotators agreed with the judgment in 80% of samples, which corresponds to 0.6 Kappa<sup>2</sup>, 20% of the corpus where there is no agreement have been removed). The rest of the body is considered a "gold standard" or "BASELINE" for evaluation.

The purpose of this maneuver is to create a body where there will be no judgments controversy.

To make their judgments and annotate the corpus annotators follow guidelines. In what follows, we will quote the various directives that were taken into account.

### 5.3.2) Directives judgments

The inference is a one-way relationship.

The assumption must be inferred from a text, but the text should not necessarily be inferred from the hypothesis.

The assumption must be inferred from the text entirely. The inference is false if there is a part of the hypothesis can not be inferred from the text.

where the inference is likely to be judged as true.

it is allowed to use the knowledge of the world as in the example *the turnover of Google is \$ 50 million*. We need to know that Google is a company so we can give it the ability to have a turnover.

### 5.3.3) The evaluation measures

The corpus annotation system adopted in the two preceding challenges is binary, that is to say the system gives two possible outcomes is the inference between the two texts is true or false}.

The result is compared to '**GOLD standard**' And the percentage giving the number of times there is similarity between the system and the '**gold standard**' gives **The 'accuracy'** system.

The accuracy is a standard measure in the systems of natural language processing. It is frequently used to evaluate the performance of applications (Beyer et al. 2005). It is calculated as follows.

**Accuracy = X / Y.**

Where:

**X** : Represents the number of times the system results are similar to the gold standard.

**Y** : Represents the number of pairs contained in the corpus.

For example the number of similar results of 500 pairs and 800 pairs of corpus, the accuracy is 500/800 which is equal to 62.5%.

### Analysis of the main methods used

In what follows, we will present the various processing steps performed to detect textual inference.

### 5.4.1) The pretreatments

Whatever the technique used to perform textual inference, it is necessary to pre-process the raw data before applying techniques inferences.

---

<sup>2</sup> Kappa (J.Cohen, 1960) is a statistical measure to calculate just how two people (or groups of individuals) A and B agree to classify N items into K mutually exclusive categories.

In the RTE three levels of pretreatments were used:  
Lexical level to avoid problems related to the morphology of words.  
Syntactic level to give a preliminary structure of the text.  
Semantic level to analyze the meaning of words.

Below we present the different existing levels of pretreatment and used for textual inference.

#### **5.4.1.1) The lexical level**

The objective of the pretreatment level of "word" is to reduce the variations due to the shape and prevent small initial errors propagate in all stages of treatment. For this, various changes have been introduced:

##### **Tokenization**

The goal of tokenization is to find the basic units of "meaning" in the texts. For this, the system must solve various problems such as managing white, punctuation, and return lines for paragraphs.

##### **Lemmatisation**

Lemmatisation a form of a word is to take its canonical form. This is defined as follows:

When it is a verb must be the meter in the infinitive:

Example:

Party (verb) -> from

In other words, they must be the masculine singular.

Example:

Party (name) -> left

To perform lexical analysis, different tool have been developed. The TreeTagger is one of the most used tools for the English language.

**The TreeTagger** tokenisation performs a stemming and labeling as shown in the following example:

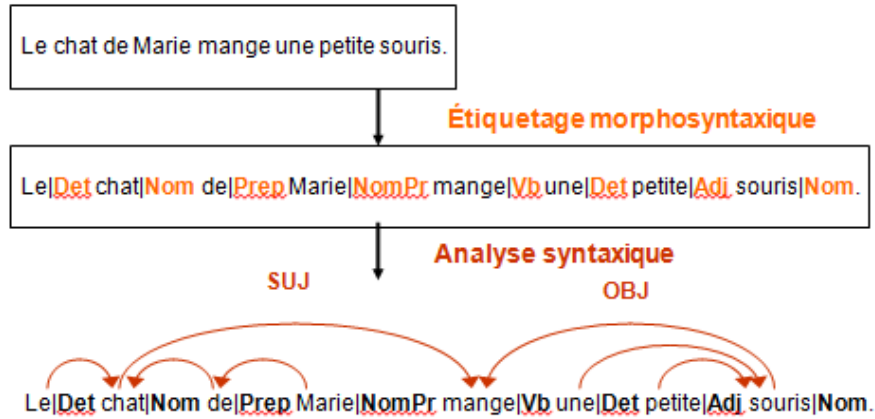
Sample entry in the **TreeTagger** "The TreeTagger easy to use."

#### **5.4.1.2) The syntactic**

The objective of this step is to describe the structures of possible sentences and analyze sentences structures.

The structure revealed by the analysis then gives precisely how the syntactic rules are combined in the text. This structure is often a hierarchy of phrases, represented by a parse tree whose nodes can be decorated (with additional information).

We illustrate this analysis with the output of one of the tools used in the syntactic annotation (Syntex)<sup>3</sup>.



**Figure 1.5: Example of syntactic annotation**

We note in the above example that the morphosyntactic analysis allows to label words and parsing can connect them.

#### 5.4.1.3) The semantic level

To simplify, we can say that the semantic analysis is based, among other things, understanding the meaning of words of text, as opposed to lexical and grammatical analysis, which analyzes the words from the lexicon and grammar. As part of the semantic analysis, it is essential to analyze the meaning of words to understand what is said. For that several approaches have been adopted to annotate relationships between words to better understand their meaning. One such approach is the predicate argument structure that is explained below.

The structure that we call predicate is a graph of predicate-argument relation, where predicates represent the action.

A predicative relationship is a relationship of dependency syntax. The predicate can have several types of arguments (subject, a direct object and indirect object complement).

Example:

<sup>3</sup> The function of this parser is to identify dependency relationships between words and extract a corpus of phrases (verbal, nominal, adjectival) (Bourigault, 2000).

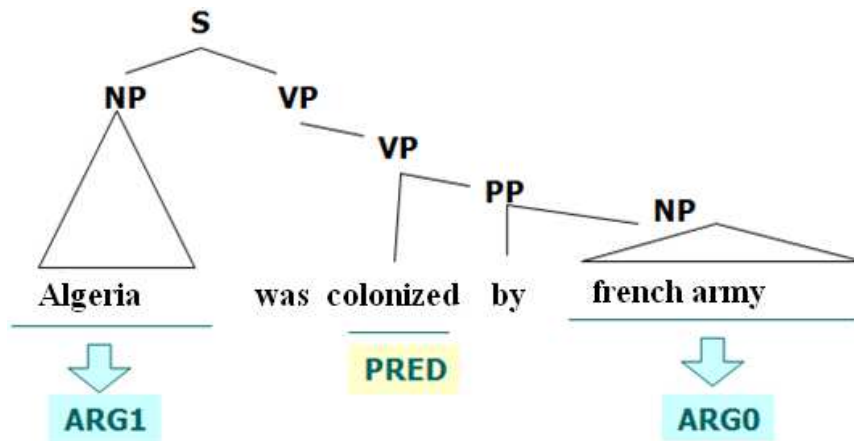


Figure 1.6: Example of predicate argument structure

#### 5.4.2) The different levels of textual inference

In this section we present the different levels of inference (Lexical, lexical and syntactic, semantic (logical) and knowledge of the world) used for the detection of textual inference.

##### 5.4.2.1) The inference at the lexical level

At this level, the inference between two text segments is accepted if there are similar words between T and H, where the words in the sentence can be inferred from H T after lexical transformations (vanderwede et al., 2005 ). The three inference techniques are below:

##### Morphological derivations

This inference mechanism considers that two terms are equivalent if one can be obtained from each other after a morphological derivation. There are three types of morphological derivations:

##### Standardization

Example:

T: "the**acquisition** an Airbus A380 by King FAHD. "

H: "King FAHD **acquired** an Airbus A380. "

The transformation <acquisition> en <acquired > allowed to deduct the inference between the two texts.

##### - The nominal bypass

Example

T: The GIA gives **dread** the Algerian people.

H: The GIA is a group **terrorist**.

The transformation of **dread** in **terrorist** allowed to deduct the inference between the two texts.

- **Relations between nouns and verbs**

Example

T: Mark wins every time.

H: Mark is a winner.

The transformation of **Mark is a winner** in **Mark wins** allowed to deduct the inference between the two texts.

**A) Ontological relationships**

An ontology is a structured set of concepts to make sense of the information. It is also a data model that represents a set of concepts in a domain and the relationships between these concepts (Bourigault, 2004). It is used to reason about the objects in this area.

The concepts are organized in a graph whose relations can be: semantic relations and subsumption relations.

The primary purpose of an ontology is to model a set of knowledge in a given field.

This inference mechanism refers to the relationship **ontological** between two terms. These relationships are listed below.

**Synonymy**

Represents a set of interchangeable words in a given context. It is often used to recognize inference.

Example

T: "Jane **shot** Mark "

H: "Jane **killed** Mark "

Another example as (" start" /" start") (" off" /" remove").

**Generalization (hypernymie)**

The Hypernymie relationship is the generic term for a class encompassing more specific instances of classes. Y is hypernym X if X is a type Y.

Example

T: "We cut the **fir** "

H: "We have cut **shaft** "

The relationship between the tree and the tree (the tree is a generalization tree) allowed inference between the two texts.

**The hyponymy**

Hyponymy relationship is the specific term used to designate a member of a class (inverse relationship Hypernymie). X is Y hyponyme if X is a type Y.

Example

T: John took a **transport means for terrestrial** to commute Toulouse paris.

H: John was in Paris Toulouse **TGV**.



The relationship between **transport means for terrestrial** and **TGV** which allowed the inference between the two texts.

#### **The relation of meronymy**

X is Y meronym if X is a part of Y.

Example:

{ } Plane as a meronym { {door}, {engine} };

#### **Knowledge of the world in the analysis lexicon**

This inference mechanism refers to the knowledge of the world to detect the inference at the lexical level (Len Schubert, 2002).

Example:

"Taliban → organization" and "yahoo → search"

#### **5.4.2.2) Inference lexical syntactic level**

The syntactic lexical level is represented by the assumption of syntactic dependency relationships.

The inference relation between T and H is defined as a collection of relations H by the relations of T, or the recovery is obtained after a sequence of transformation applied to the relationship of T. S different types of transformations are specified by:

#### **Syntactic transformations**

In this inference mechanism, the transformation is between the syntactic structures that have the same lexical items and preserve the sense of the relationship between them (Vanderwende et al .., 2005).

This kind of mechanism includes active and passive transformation affixing<sup>4</sup>.

Example:

"My cat, this sweet little Siamese, is sitting on the table." "It can become: My cat is sitting on the table, this sweet little Siamese! "

#### **The inference based on paraphrases**

In this inference mechanism, the transformation changes the syntactic structure of the text segment and some lexical items, but it keeps the inference relation between the segment and the original text is changed.

This type of relationship between the two segments is called in the literature "Paraphrase." Methods to perform the transformation proposed in (Lin and Pantel, 2001).

---

<sup>4</sup> The affixing is a grammatical construction in which two elements, normally noun phrases, are placed beside each other, with an element to define or modify the other .. When this device is used, the two are supposed to be in apposition. For example, in the phrase "my friend Alice" the name "Alice" is in apposition to "my friend."

Example:

T: "This drug is sold only in Canada."

H: "The marketing of this drug is made in Canada only."

### **Coreference**

Relationship Coreference connects a pronoun and a remote antecedent one another in the sentence. For example:

"**Italy and Germany** both played two games **they** have lost a game yet. "

### **Infers**

"Neither Italy nor Germany has not yet lost a game," this includes the transformation of coreference "**They** → **Italy and Germany.** "

### **5.4.2.3) The semantic inference (logic)**

At this level, the inference between two text segments is accepted if the meaning of the two sentences is consistent. In other words, the textual inference is regarded as a problem of logical implication between the meaning of the two sentences (Tatu et al., 2006).

For this, the predicate argument structure is often used, that is to say, the text segments T and H are transformed into predicate and through logical deductions such as the use of (rebuttal evidence<sup>5</sup>) You get to deduct the inference.

An example of systems using this inference method described in section (5.5.4.2).

### **5.4.3) The resources used**

In the various techniques textual inference more resources are used (WordNet fram-net, Cyc ...). The set is a complete "ecosystem" covering lexical, syntactic and semantic aspects. Together, these resources provide a good starting point for semantic developments in TAL or in the Semantic Web, such as information retrieval, inference for automatic reading comprehension, word sense disambiguation, the resolution of anaphora and also textual inference. In what follows, we will define the various existing resources and used to detect textual inference.

### **5.4.3.1) The WordNet**

WordNet (Miller, 1995) is a lexical database developed since 1985 by linguists of Cognitive Science Laboratory at Princeton University. It is a semantic network of the English language, which is based on a psychological theory of language. The first version released back in June 1991. Its purpose is to identify and classify relate in various ways the semantic and lexical content of the English language. The system is in the form of an electronic database (Chaumartin, 2007).

The synset (set of synonyms) is the atomic component which is based on WordNet. A synset is a group of words denoting direction or a particular purpose. A synset is defined by its relationship with neighboring sense. The nouns and verbs are organized

---

<sup>5</sup> The refutation is a logical process of proving the falsity or failure of a proposal or an argument.

into hierarchies. Hyperonymy of hyponymy relations and connect the "ancestors" of nouns and verbs with their "specialties". At the root level, these hierarchies are organized into basic types.

Like a traditional dictionary, WordNet offers and, for each word, a list of synsets corresponding to all its meanings listed. Synsets but also have other uses: they can represent more abstract concepts, higher-level words and their meaning, which can hold as ontology. We can query the system about the hyperonyms of a particular word. As example of the most common meaning of the name "car" (corresponding to the synset "1. Because self ..."), the relationship defines a hyperonymy tree concepts more general:

- 1. car, auto, automobile, machine, motorcar
  - => Motor vehicle, automotive vehicle
    - => Vehicle
      - => Conveyance, transportation
        - => Instrumentality, instrumentation
          - => Artifact, artefact
            - => Object, physical object
              - => Entity, something

In this example, it is clear that the latter concept, "entity, something" is more general, more abstract (and it could be the super-concept in a variety of more specialized concepts).

We can also query the system about the inverse relationship of hypernymie, thehyponymy. WordNet actually offers a multitude of other ontologies, making use of more specialized and restrictive semantic relations. We can query the system about the meronyms a word or a concept, the constituent parts of an object ("HAS-PART"). The meronyms associated sense "because self ..." the word "car" are:

- 1. car, auto, automobile, machine, motorcar
  - HAS PART: accelerator, accelerator pedal, gas pedal, gas, throttle, gun
  - HAS PART: air bag
  - HAS PART: auto accessory
  - HAS PART: automobile engine
  - HAS PART: automobile horn, car horn, motor horn, horn
  - (...)

#### **5.4.3.2) The FrameNet**

FrameNet (Baker, Fillmore, and Lowe, 1998), led to Berkeley at the initiative of Charles Fillmore project is based on the semantics of frames (frame semantics). Documenting the syntactic and semantic combinatorial for each direction of a lexical entry through a manual annotation selected in examples on corpus aims FrameNet representativeness criteria lexicographical. The annotations are then summarized in tables that summarize each word frames with their syntactic arguments.

#### **5.4.3.3) The Cyc**

Cyc is an artificial intelligence project launched in 1984 by Doug Lenat. Cyc is to combine an ontology and a database full of common sense data, to allow applications of artificial intelligence application. To conduct similar to human reasoning. Frag-

ments typical knowledge eg "cats have four legs" " *Paris is the capital of the France* ". They contain terms (PARIS, FRANCE, CAT?) And assertions (" *Paris is the capital of the France* ") That connect these words together. Thanks to the inference engine provided with the base Cyc, it is possible to get an answer to a question like "What is the capital of the France ? " The Cyc database contains millions of assertions (facts and rules) cash in hand.

### 5.1) The analysis of systems participating in RTE 2

We marked each research group participating in RTE 2 types of inference used. The results are displayed in Table 1.6.

**6. Table 1.1 Representation of different types of inferences made by research groups**

Type of analysis and Research Groups	lexical	syntactic	lexical-semantic	Logic	Digital	Time
<b>UNED</b>	+	+	+		+	
<b>Umess</b>	+					
<b>MITRE</b>	+			+		
<b>TSRI</b>	+	+				
<b>GOGEX</b>	+		+	+		
<b>LCC'S</b>	+		+			
<b>C &amp; C</b>	+	+				

### 5.5.4) Some examples of inference used by research groups

In the RTE 2 we noticed that all research groups did not use temporal inference in their systems and present the results of RTE 3 are not officially released yet, but according to our reading various publications of research groups participating in RTE3, there are two groups who have alluded to the temporal inference. For this we have chosen to describe their systems.

#### 5.5.4.1) The recognition of textual inference based on dependency analysis and WordNet (National University of Distance Education in Madrid)

The system presented shows how semantic information can be extracted from text using syntactic structuring data dependency analysis, and lexical-semantic resources as Word Net can develop RTE.

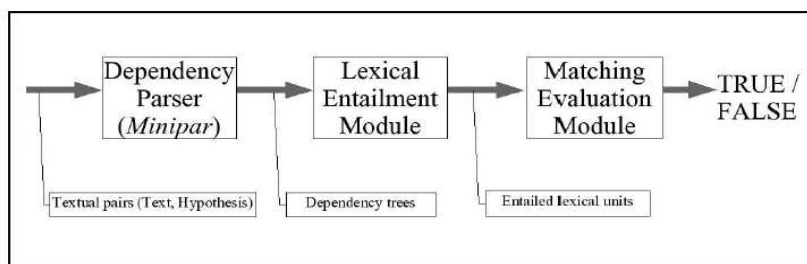
The techniques used by this system are as follows:

- the length of the text and the hypothesis analysis.
- lexical inference between nodes of trees using Word Net.

- the correlation between the dependency trees based on the notion of inclusion.

#### A) The system architecture

The system architecture is shown in the following figure (Figure 1.7):



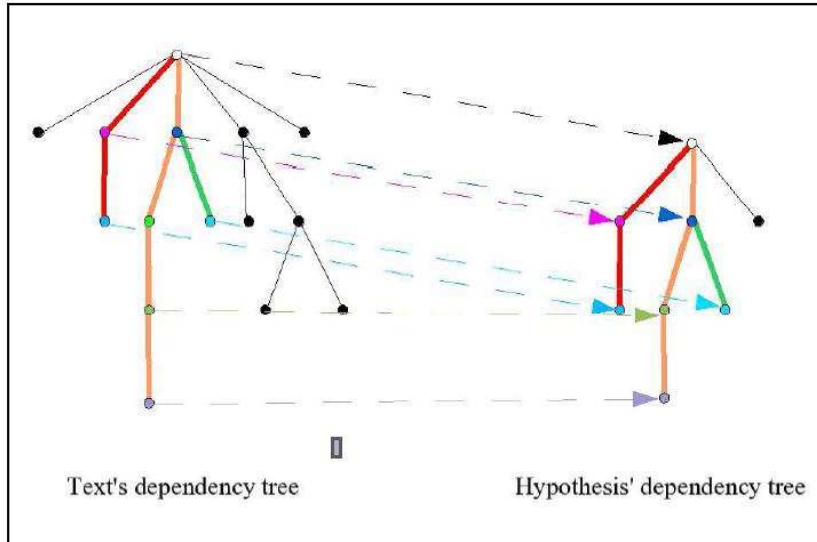
**Figure 1.7: The system architecture**

This architecture consists of three modules:

**The dependency analysis:** It is to standardize the information in the dataset, generate existing between words and give the output a dependency tree consists of nodes that represent the words of the sentence and arcs that represent dependencies between nodes dependencies. This work is done by a software named "Lin's Minipar."

**Lexical analysis:** takes the information from the dependency analysis and returns the words of the hypothesis H that are inferred from the text T. This module uses Word-Net to detect relations (synonymy, hyponymy, meronymy) between lexical units.

**The relationship between dependency trees :** The goal is to deduce whether the tree assumption is covered by the dependency tree of the text, for this, the rule is that arc is covered if it is said in the same location as in tree representing the text and there is an inference between these nodes and the text. The figure below (Figure 1.8) shows this kind of recovery.



**Figure 1.8: Example of overlap between dependency tree**

**B) The experimental system**

The group soumi two systems challenge.

- **System 1**

The system1 uses only the first two modules and the decision on the existence of inference is taken relative to the number of nodes in the hypothesis inferred from the dependency tree of the text.

- **System 2**

The system uses two 3 modules and the decision is made in relation to the number of arcs covered.

The results are displayed in the **Table 1.2**. Using WordNet only gave good results, but adding the module cover it decreases system performance.

Systems	Accuracy
System 1:	56.37%
System 2:	54.75%

**Table 1.2: Accuracy values systems**

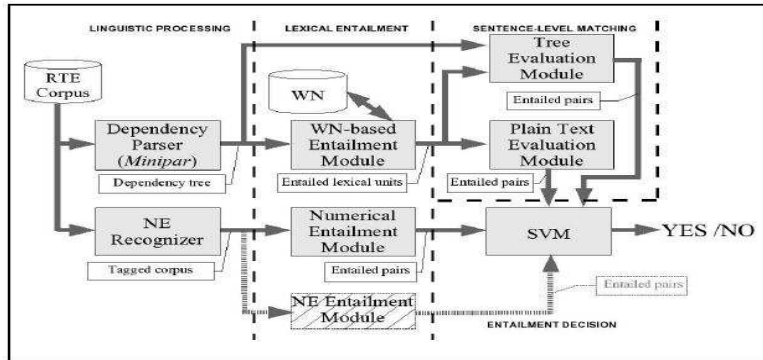
The concept of recovery is not appropriate for the RTE as a wide collection does not imply a semantic inference, and low recovery does not imply a semantic difference. Using Word Net has contributed to the inference at the lexical level and increased system performance. In this direction, the next steps will be to identify and evaluate inferences from numerical expressions, **named entities**<sup>6</sup> and temporal expressions.

---

<sup>6</sup> Named entities means all names of people, places, business contained in a text.

### C) The evolution of the system

Which was developed for RTE 2 is a module for the detection of numerical expressions, which has increased significantly **precision** (Harrera et al., 2006). The following figure shows how the module is introduced into their system.



**Figure 1.9: System Architecture UNED**

In the RTE 3, the group focused on inference between the named entities. He defined the inference relations between named entities (Rodrigo et al., 2007). Example:

- Own name infers E1 E2 own name if a string contains the E1 E2 chain.
- an expression of the time t1 infers an expression time T2 if the time t1 is expressed in included in the interval T2.

This inference module has also contributed to increase the accuracy (Rodrigo et al, 2007).

#### 5.5.4.2) COGEX (University of Texas, USA)

The system uses a logical approach to solving the textual inference. In other words, the textual inference is regarded as a problem of logical implication between the meaning of the two sentences (Tatu et al., 2006).

The description of the system and the changes that occurred in each challenge is described in the following.

### A) The system description

The first step is to transform the text and the hypothesis logical form (Moldovan and Rus, 2001).

To do this we must first transform the nature of language has a predicate size argument for the cella group uses WordNet to link with the predicate argument. Specifically WordNet produces relations between synsets, and each synset corresponding predicate.

The predicate may have one or more arguments and the predicate is the name of a single argument in general, and the predicate is a verb has three arguments: the event, the subject and the object complement.

For each relationship in the lexical chain<sup>7</sup>The system generates an axiom using predicates corresponding synset in the relationship.

For example: there is an inference relation between the verb **sale** and the verb **pay**.

The system generates the following axiom for this relationship:

Vendre\_VB\_1 (e1, x1, x2)  $\rightarrow$  payer\_VB\_1 (e1, x1, x3)

This type of axiom contributes to the inference when a lexical chain is found.

After the transformation of the two pairs of text in logical format the group uses evidence by "absurd" or "proof by contradiction" (Wos 1998). The negation of the hypothesis H is performed if there is a conflict or contradiction of a deduction from the text T, we conclude that the hyponym is differentiable text.

### B) The evolution of the system

It was developed for RTE 2 module that treats negation in transformation the text predicate and another module that is a semantic analysis as a pretreatment to give the existing relationship between the verb and its arguments as well as between arguments themselves (Tatu et al., 2006).

For RTE3 the group has developed and integrate their system has several tools. In what follows we present the system architecture and new tools developed and used to improve the inference. The diagram of the system for the last three TEN by the group is given by the figure below.

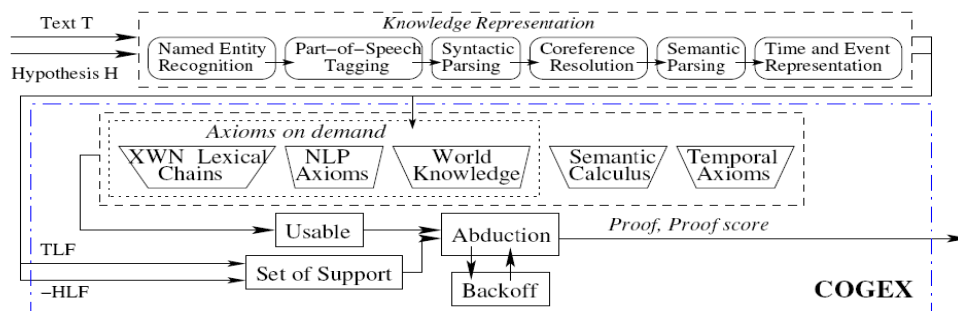


Figure 1.10: System Architecture

#### - EXtended WordNet

XWN (eXtended WordNet) is a project that aims to enrich relations WordNet dictionary with semantic relations between synsets and turn them into logical format (Tatu and Moldovan, 2007).

#### - TARSQI

It is a modular system for automatic temporal annotation that adds time expressions, events and temporal relations to news texts (Venhaguane et al., 2005).

<sup>7</sup> A lexical chain is a chain where there is a relationship between two synsets.



#### - **Tool for managing coreferences**

To link sentences in longer texts, and solve the problem that is brought by the coreference in textual inference, the tool developed algorithm combines Hobbs (Hobbs, 1978) and the algorithm of anaphora resolution (Lappin and Leas 1994). For the RTE, it is important to have the relationship between the predicates of a long text.

Example 1: *George Bush grew up in Greenwich, Connecticut, is then a member of a secret brotherhood became famous student.*

*Link George Bush and there*, is one of the tasks that the tool should solve.

Development XWN-KB has had a considerable impact on the TEN, but the use of TARSQI yielded no impact on the result as the use of temporal expressions in the corpus is non-existent.

In the work begun by the UNED named entities, the group has established several rules of inference between named entities, among which is an inference rule between temporal expressions. This can be seen as an implicit contribution to the temporal inference. But realistically the temporal inference is considered a prospect for their future research.

## **5 Conclusion**

In this paper we explored the contribution of the TEN in different NLP applications (IR, QR, AR and RA) and we explored different approaches to detect inference (lexical, syntactic lexical, semantic and logical). As we analyzed the different approaches research groups that participated in the challenge Pascal RTE. This step allowed us to discover the paths that have not yet been taken to detect textual inference.

Finally we focused to describe systems that have mentioned the temporal aspect in their research. We noticed that in the three RTE that occurred, the temporal inference is a perspective that has not yet started. We will precisely describe in the next chapter the temporal aspect in the RTE.

## **References**

1. D. Bouneffouf. Role of temporal inference in the recognition of textual inference, arXiv preprint arXiv: 1302.5645, 2013
2. Baker, Charles J. Fillmore, and John B. Lowe. 1998. The Berkeley FrameNet project. In Proceedings of the COLING-ACL, Montreal.
3. Emile Benveniste, Problems in General Linguistics. Paris, Gallimard, vol. II.
4. BOURIGAULT D. Recent Advances in Computational Terminology, 2000.
5. BOURIGAULT D. AUSSENAC-GILLES N. Charlet and J. Construction terminology or ontological resources from texts: a unifying framework for three case studies, Journal of Artificial Intelligence, 18 (4), 24 pp., 2004.
6. Charolles Mr. "L" discourse framing - universe, fields, areas and spaces, "Cahier de recherche linguistics, 6, p. 1-73. 1997.

7. Francois-Regis Chaumartin, wordnet and its ecosystem, BDL-CA, 2007, montreal.
8. Cohen J. "A coefficient of agreement for nominal scales," *Educ. Psychol. Meas.*: 20, 27-46. 1960
9. Textual inference problems from the PASCAL RTE. Challenge, 2005.
10. Len Schubert. Can we derive general Word Knowledge from Texts?. 2002.
11. Part-of-Speech Tagging with Neural Networks. Proceedings of the 15th International Conference on Computational Linguistics (COLING-94). August 1994.
12. T. Joachims, Information Retrieval and Language Technology (pdf), 2003, Cornell University.
13. Leila Kosseim, bilingual information extraction, 2005.
14. Gerard Ligauzat. Knowledge representation and linguistics. Armand Colin, Paris, 1994.
15. Dekang Lin and Patrick Pantel. Discovery of inference rules for Question Answering. Natural Language Engineering, 2001
16. Automatic translation. France. North University Press, 1996. p. 15-16.
17. C.Macleod, R.Grishman, A.Meyers, L.Barrett and R. Reeves. 1998. Nomex: A lexicom of normalisations.in 8 Proceedings of the International Congress of the European Association for lexicography.1998. liege, begium: EURALEX.
18. Mani and George Wilson .. Processing of News. In Proceedings of the 38th Annual Meeting of the Association for Computational Linguistics (ACL2000), pages 69-76, 2000
19. Dan I. Moldovan and Vasile Rus. Logic form transformation of wordnet and Its applicability to answering questions. In Meeting of the Association for Computational Linguistics, pages 394-401, 2001.
20. Moldovan and Rus Logic Forms Can Be Utilized by a wide variety.2001.
21. P. Miller. "Notes on phonology and orthography in several Katuic Mon-Khmer groups in Northeast Thailand." *Mon-Khmer Studies* 24: 27-51, 1995.
22. Pierre Nugues. An Introduction to Language Processing with Perl and Prolog. Springer Verlag, 2006.
23. E. Nyberg, T.Mitamura, J. Carbonell, J. Callan, K. Clins-Thompson, K Czuba, Duggan, L. Hiyakumoto, N. Hu, Y. Huang, J. Ko, L.V. Lita S.Muratagh and V. Pedro. The JAVELIN Question-Answering System at TREC 2002. In proceeding of the 11th Text Retrieval Conference (TREC-11), 2002.
24. Paul Kiparsky and Carol Kiparsky. In Manfred Bierwisch and Karl Erich Heidolph, editors, *Progress in Linguistics. A Collection of Papers*, pages 143-173. Mouton, Paris.
25. A. Rodrigo, A. Pe ~ nas, J. Herrera and F. Verdejo .. The Effect of Entity Recognition on Answer Validation. In *Lecture Notes in Computer Science*. In press 2007.
26. Daniel Sleator and Davy Temperley. 1991. Parsing English with a Link Grammar. Carnegie Mellon University Computer Science technical report CMU-CS-91-196 October 1991.
27. Gérard Swinnen Learning to program with Python, Copyright 2005.
28. Marta Tatu and Dan Moldovan. Recognising COGEX at the third of textual entailment challenge. In *procédure of the wokshop on textual entailment*, Prague, June 2007.
29. Marta Tatu, B Iles, J. Slavick, A. Novischi, and D. Moldovan., COGEX at the third of Recognising textual entailment challenge. In *procédure of the wokshop on textual entailment*, Venice, Italy, 2006
30. Lucy vanderwende, Deborah Coughlin and Bill Dolan.
31. Contribute 2005.what syntax in entailment task. In *proceedings of pascalchallenge workshop on recogning texual entailment*, 2005.
32. Mr. Venhagen, I. Mani, R. Sauri, R. Knippen, J. Littman and J. Pustejovsky .. Automating Tenporal Annotation With TARSQI. In *Proceedings of ACL 2005*. demo session.
33. L. WOS. Automated Reasoning -33 Basic Research Problems. Prentice-Hall. 2005

34. François Yvon. A small introduction to Automatic Natural Language Processing,. Conference on Knowledge Discovery and Data Mining: p. 27-36. ACM, New York, 2010.
35. Sohn, T., Li, K. A., Griswold, W. G. Hollan, J. D.: A Diary Study of Mobile Information Needs. Proceedings of the twenty-sixth annual SIGCHI conference on Human factors in computing, p. 433-442, ACM, Florence, 2008.
36. Wu, Z. and Palmer, M.: Verb Semantics and Lexical Selection. In Proceedings of the 32nd Meeting of the Association for Computational Linguistics, pp.133-138 (1994).