Numerical Data Integration for Cooperative Question-Answering

Véronique Moriceau

Institut de Recherche en Informatique de Toulouse

118, route de Narbonne

31062 Toulouse cedex 09, France

`moriceau@irit.fr`

Abstract

In this paper, we present an approach which aims at providing numerical answers in a question-answering system. These answers are generated in a cooperative way: they explain the variation of numerical values when several values, apparently incoherent, are extracted from the web as possible answers to a question.

1 Introduction

Search engines on the web and most existing question-answering systems provide the user with a set of hyperlinks and/or web page extracts containing answer(s) to a question. These answers may be incoherent to a certain degree: they may be equivalent, complementary, contradictory, at different levels of precision or specifity, etc. It is then quite difficult for the user to know which answer is the correct one.

In a cooperative perspective, we propose an approach for answer generation in natural language which uses answer **integration**. When several possible answers are selected by the extraction engine, the goal is to define a coherent core from candidate answers and to generate a **cooperative answer**, i.e. an answer with explanations. We assume that all web pages are equally reliable since page *provenance* information (defined in (McGuinness and Pinheiro da Silva, 2004) e.g., source, date, author, etc.) is difficult to obtain.

To adequately deal with data integration in question-answering, it is essential to define precisely relations existing between potential answers. In this introduction, we first present related works. Then, we define a general typology of relations between candidate answers.

1.1 Related works

Most of existing systems on the web produce a set of answers to a question in the form of hyperlinks or page extracts, ranked according to a relevance score. For example, COGEX (Moldovan et al., 2003) uses its logic prover to extract lexical relationships between the question and its candidate answers. The answers are then ranked based on their proof scores. Other systems define relationships between web page extracts or texts containing possible answers: for example, (Radev and McKeown, 1998) and (Harabagiu and Lacatusu, 2004) define *agreement* (when two sources report the same information), *addition* (when a second source reports additional information), *contradiction* (when two sources report conflicting information), etc. These relations can be classified into the 4 relations defined by (Webber et al., 2002), i.e. **inclusion**, **equivalence**, **aggregation** and **alternative** which we present below.

Most question-answering systems provide answers which take into account neither information given by all candidate answers nor their inconsistency. This is the point we focus on in the following section.

1.2 A general typology of integration mechanisms

To better characterize our problem, we collected a corpus of about 100 question-answer pairs in French that reflect different inconsistency problems (most of pairs are obtained via Google or QRISTAL[1]). We first assume that all candidate answers obtained via an extraction engine are potentially correct, i.e. they are of the semantic type expected by the question.

---

[1] www.qristal.fr, Synapse Développement.

For each question of our corpus, a set of possible answers is extracted from the web. The goal of our corpus analysis is to identify relations between those answers and to define a general typology of associated integration mechanisms. We use for this purpose the 4 relations defined in (Webber et al., 2002) and for each relation, we propose one or several integration mechanisms in order to generate answers which take into account characteristics and particularities of candidate answers.

1.2.1 Inclusion

A candidate answer is in an inclusion relation if it entails another answer (for example, concepts of candidate answers linked in an ontology by the *is-a* or *part-of* relations). For example, *in Brittany* and *in France* are correct answers to the question *Where is Brest?*, linked by an inclusion relation since *Brittany* is a part of *France*.

1.2.2 Equivalence

Candidate answers which are linked by an equivalence relation are consistent and entail mutually. The corpus analysis allows us to identify two main types of equivalence:

(1) Lexical equivalence: use of acronyms or foreign language, synonymies, metonymies, paraphrases, proportional series. For example, to the question *Who killed John Lennon?*, *Mark Chapman, the murderer of John Lennon* and *John Lennon's killer Mark Chapman* are equivalent answers.

(2) Equivalence with inference: in a number of cases, some common knowledge, inferences or computations are necessary to detect equivalence relations. For example, *The A320 is 22* and *The A320 was built in 1984* are equivalent answers to the question *How old is the Airbus A320?*.

1.2.3 Aggregation

The aggregation relation defines a set of consistent answers when the question accepts several different ones. In this case, all candidate answers are potentially correct and can be integrated in the form of a conjunction of all these answers. For example, an answer to the question *Where is Disneyland?* can be *in Tokyo, Paris, Hong-Kong and Los Angeles*.

1.2.4 Alternative

The alternative relation defines a set of inconsistent answers. In the case of questions expecting a unique answer, only one answer among candidates

is correct. On the contrary, all candidates can be correct answers.

(1) A simple solution is to propose a disjunction of candidate answers. For example, if the question *When does autumn begin?* has the candidate answers *Autumn begins on September 21st* and *Autumn begins on September 20th*, an answer such as *Autumn begins on either September 20th or September 21st* can be proposed. (Moriceau, 2005) proposes an integration method for answers of type *date*.

(2) If candidate answers have common characteristics, it is possible to integrate them according to these characteristics ("greatest common denominator"). For example, the question *When does the "fête de la musique" take place?* has the following answers *June 1st 1982, June 21st 1983, ..., June 21st 2005*. Here, the extraction engine selects pages containing the dates of music festivals over the years. Since these candidate answers have day and month in common, an answer such as *The "fête de la musique" takes place every June 21st* can be proposed (Moriceau, 2005).

(3) Numerical values can be integrated in the form of an interval, average or comparison. For example, if the question *How far is Paris from Toulouse?* has the candidate answers *713 km, 678 km* and *681 km*, answers such as *Paris is at about 690 km from Toulouse* (average) or *The distance between Paris and Toulouse is between 678 and 713 km* (interval) can be proposed.

2 Motivations

In this paper, we focus on answer elaboration from several answers of type *numerical* (case (3) above). Numerical questions deal with numerical properties such as distance, quantity, weight, age, etc. In order to identify the different problems, let us consider the following example :
What is the average age of marriage in France?
- In 1972, the average age of marriage was 24.5 for men and 22.4 for women. In 2005, it is 30 for men and 28 for women.
- According to an investigation carried out by FNAIM in 1999, the average age of marriage is 27.7 for women and 28.9 for men.
- The average age of marriage in France increased from 24.5 to 26.9 for women and from 26.5 to 29 for men between 1986 and 1995.

This set of potential answers may seem incoherent but their internal coherence can be made ap-

parent once a variation criterion is identified. In a cooperative perspective, an answer can be for example:

In 2005, the average age of marriage in France is 30 for men and 28 for women.
It increased by about 5.5 years between 1972 and 2005.

This answer is composed of:

1. a direct answer to the question,
2. an explanation characterizing the variation mode of the numerical value.

To generate this kind of answer, it is necessary (1) to integrate candidate answers in order to elaborate a direct answer (for example by solving inconsistencies), and (2) to integrate candidate answers characteristics in order to generate an explanation.

In the following sections, we first define a typology of numerical answers and then briefly present the general architecture of the system which generates cooperative numerical answers.

2.1 A typology of numerical answers

To define the different types of numerical answers, we collected a set of 80 question-answer pairs about prices, quantities, age, time, weight, temperature, speed and distance. The goal is to identify for each question-answer pair:
- if the question expects one or several answers (learnt from texts and candidate answers),
- why extracted numerical values are different (is this an inconsistency? an evolution?).

2.1.1 The question accepts only one answer

For example, *How long is the Cannes International Film Festival?*. In this case, if there are several candidate answers, there is an inconsistency which has to be solved (cf. section 4.1).

2.1.2 The question accepts several answers

This is the case when numerical values vary according to certain criteria. Let us consider the following examples.
Example 1 :
How many inhabitants are there in France?
- Population census in France (1999): 61632485.
- 61.7: number of inhabitants in France in 2004.
In this example, the numerical value (quantity) is a property which changes over time (*1999, 2004*).

Example 2 :
What is the average age of marriage of women in 2004?

- In Iran, the average age of marriage of women went from 19 to 21 years in 2004.
- In 2004, Moroccan women get married at the age of 27.
In this example, the numerical value (age of marriage) varies according to place (*in Iran, Moroccan*).

Example 3 :
At which temperature do I have to serve wine?
- Red wine must be served at room temperature.
- Champagne: between 8 and 10 ˚C.
- White wine: between 8 and 11 ˚C.
Here, the numerical value (temperature) varies according to the question focus (wine).

The corpus analysis allows us to identify 3 main variation criteria, namely *time, place* and *restriction* (restriction on the focus, for example: Champagne/wine). These criteria can be combined: some numerical values vary according to time and place, to time and restrictions, etc. (for example, the average age of marriage vary according to time, place and restrictions on men/women). Note that there are several levels of restrictions and that only restrictions of the same type can be compared (cf. section 3.2). For example, *metropolitan population* and *population of overseas regions* are restrictions of the same ontological type (geographical place) whereas *prison population* is a restriction of a different type and is not comparable to the previous ones.

2.2 Architecture of the system

Figure 1 presents the general architecture of our system which allows us to generate answers and explanations from several different numerical answers.

Questions are submitted in natural language to QRISTAL which analyses them (focus, answer expected type) and which selects potential answers from the web: QRISTAL searches web pages containing the keywords of the query and synonyms (extraction engine). Then, an extraction grammar constructs a set of frames from candidate web pages. From the frame set, the variation criteria and mode of the searched numerical value are identified. Finally, a natural language answer is generated explaining those characteristics. Each of these stages is presented in the next sections.
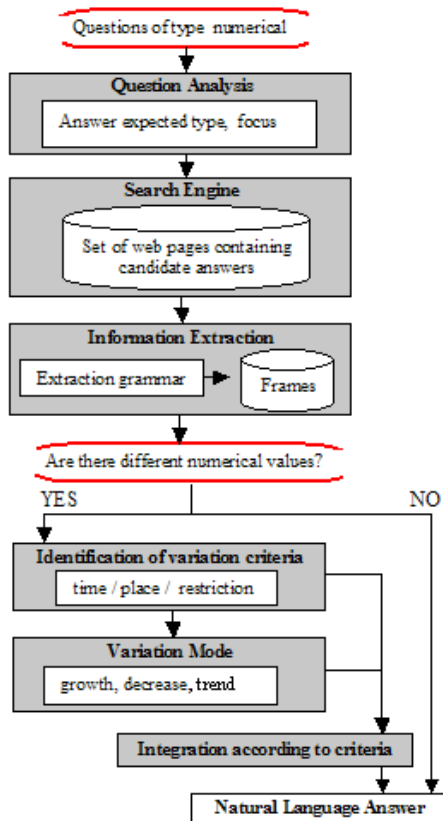
Figure 1: Architecture

3 Answer characterization

Answer characterization consists in 2 main stages:
- information extraction from candidate web pages,
- characterization of variation (criteria and mode) of numerical values if necessary.

3.1 Answer extraction

Once QRISTAL has selected candidate web pages (those containing the question focus and having the expected semantic type), a grammar is applied to extract information needed for the generation of an appropriate cooperative answer. This information is mainly:
- the searched numerical value (*val*),
- the *unit* of measure,
- the question focus and its synonyms (*focus*) (for the moment, synonyms are not considered because it requires a lot of resources, especially in an open domain),
- the *date* and *place* of the information,
- the *restriction(s)* on the question focus (essentially, adjectives or relative clauses).

In addition, the corpus analysis shows that some other information is essential for infering the precision degree or the variation mode of values. It is mainly linguistic clues indicating:
- the *precision* of the numerical value (for example adverbs or prepositions such as in *about 700, ...*),
- a *variation* of the value (for example temporal adverbs, verbs of change/movement as in *the price* **increased** *to 200 euro*).

We define a frame a_i which gathers all this information for a numerical value:

$$a_i = \begin{bmatrix} Val = \\ Precision = \\ Unit = \\ Focus = \\ Date = \\ Place = \\ Restriction = \\ Variation = \end{bmatrix}$$

A dedicated grammar extracts this information from candidate web pages and produces the set A of N candidate answers: $A = \{a_1, ..., a_N\}$. We use a gapping grammar (Dahl and Abramson, 1984) to gap elements which are not useful (in the example **According to an investigation carried out by FNAIM** *in 1999, the average age of marriage is 27.7 for women and 28.9 for men*, elements in bold are not useful). We give below the main rules of the grammar, optional elements are between brackets:

Answer → Nominal Sentence | Verbal Sentence
Nominal Sentence → Focus (Restriction), ..., (Date), ..., (Place), ..., (Precision) Val (Unit)
Verbal Sentence → Focus (Restriction), ..., (Date), ..., (Place), ..., Verb, ..., (Precision) Val (Unit)
Verb → VerbQuestion | Variation
VerbQuestion → *count* | *estimate* | *weigh* | ...
Variation → *go up* | *decrease* | ...
Precision → *about* | *on average* | ...
Place → Country | City | ...
Time → Date | Period | ...
Restriction → Adjective | Relative | ...
....

Figure 2 shows an extraction result.

A syntactic analysis is also necessary to check the relevance of extracted information. For example, suppose that the answer *population of cities of France which have more than 2000 inhabitants (...)* is proposed to the question *How many inhabitants are there in France?*. A syntactic analysis has to identify the expression *cities of France which (...)* as a restriction of *population* and to

```
What is the price of a Peugeot 206?
(1) Ads (August 2005): sell Peugeot 206 diesel, 3400 km,
    16100 €, City: Toulouse, Tel: ...
(2) December 2005:
    Peugeot 206 (gas): 17200 €, 2700 km, Address: ... Paris
```

$$a_1 \begin{pmatrix} Val = 16100 \\ Precision = \varnothing \\ Unit = € \\ Focus = \text{Peugeot 206} \\ Date = \text{August 2005} \\ Place = \text{Toulouse} \\ Restriction = \text{diesel,} \\ \qquad\qquad 3400\ km \\ Variation = \varnothing \end{pmatrix} \quad a_2 \begin{pmatrix} Val = 17200 \\ Precision = \varnothing \\ Unit = € \\ Focus = \text{Peugeot 206} \\ Date = \text{December 2005} \\ Place = \text{Paris} \\ Restriction = \text{gas,} \\ \qquad\qquad 2700\ km \\ Variation = \varnothing \end{pmatrix}$$

Figure 2: Results of extraction

infer that *2000 inhabitants* is a property of those cities and that it is not an answer to the question. We plan to investigate the lexical elements (prepositions, predicative terms, etc.) necessary to this analysis.

Elements of extraction evaluation are presented in figure 3: we submitted 30 questions to Google and QRISTAL. Our system can select the correct direct answer provided that QRISTAL returns the correct answer among relevant pages (for 87% of the questions we evaluated) and that our grammar succeeds in extracting relevant information (this has to be evaluated).

Google		
Correct Answer's Rank (average)	Incorrect Answer before the correct one	Relevant Pages (first 30 links)
4	43%	23%
QRISTAL		
Correct Answer	Correct Answer among relevant pages	Relevant Pages (first 30 links)
13%	87%	30%

Figure 3: Elements of evaluation

3.2 Variation criteria

Once we have the frames representing the different numerical values, the goal is to determine if there is a variation and to identify the variation criteria of the value. In fact, we assume that there is a variation if there is at least k different numerical values with different criteria (time, place, restriction) among the N frames (k is a rate which depends on N: the more candidate answers there are, the greater is k). Thus, a numerical value varies according to:

1. **time** if $card\,(\{a_i,\ such\ as\ \exists\,a_i,\ a_j \in A,$
    $a_i(Val) \neq a_j(Val)$
    $\wedge\quad a_i(Unit) = a_j(Unit)$
    $\wedge\quad a_i(Date) \neq a_j(Date)\ \}) \geq k$

2. **place** if $card\,(\{a_i,\ such\ as\ \exists\,a_i,\ a_j \in A,$
    $a_i(Val) \neq a_j(Val)$
    $\wedge\quad a_i(Unit) = a_j(Unit)$
    $\wedge\quad a_i(Place) \neq a_j(Place)\ \}) \geq k$

3. **restriction** if $card\,(\{a_i,\ such\ as$
    $\exists\,a_i,\ a_j \in A,\quad a_i(Val) \neq a_j(Val)$
    $\wedge\quad a_i(Unit) = a_j(Unit)$
    $\wedge\, a_i(Restriction) \neq a_j(Restriction)\})$
    $\geq k$

4. **time and place** if $(1) \wedge (2)$

5. **time and restriction** if $(1) \wedge (3)$

6. **place and restriction** if $(2) \wedge (3)$

7. **time, place and restriction** if $(1) \wedge (2) \wedge (3)$

Numerical values can be compared only if they have the same unit of measure. If not, they have to be converted.

For each criterion (time, place or restriction), only information of the same semantic type and of the same ontological level can be compared. For example, *metropolitan population* and *prison population* are restrictions of a different ontological level and cannot be compared. In the same way, place criteria can only be compared if they have the same ontological level: for example, prices *in Paris* and *in Toulouse* can be compared because the ontological level of both places is *city*. On the contrary, prices *in Paris* and *in France* cannot be compared since the ontological levels are respectively *city* and *country*. Several ontologies of geographical places exist, for example (Maurel and Piton, 1999) but a deep analysis of restrictions is necessary to identify the kind of implied knowledge.

In the particular cases where no information has been extracted for some criteria, it is also necessary to define some comparison rules. Thus, let $crit \in \{time,\ place,\ restriction\}$ and $a_i,\ a_j \in A$,
- if no information has been extracted for 2 compared criteria, then we consider that those criteria are equal (there is no information indicating that there is a variation according to those criteria), i.e. if $a_i(crit) = \emptyset$ and $a_j(crit) = \emptyset$, then

$$a_i(crit) = a_j(crit)$$
- if no information has been extracted for one of the 2 compared criteria, then we consider that those criteria are different (there is a variation), i.e. if $a_i(crit) = \emptyset$ and $a_j(crit) \neq \emptyset$, then $a_i(crit) \neq a_j(crit)$

In the example of figure 2, the price varies according to time, place and restriction. In the following example (figure 4), the price of gas varies according to time (*September 2005/ $\emptyset$*) and place (*Paris/Toulouse*). For space reasons, we give only 2 frames for each example but it is obviously not sufficient to conclude.

What is the price of 1L of gas?
(1) In September 2005, 1L of gas costs 1.51 € in Paris.
(2) Price of gas (Toulouse): 1L, 1.29 €

$$a_1 \begin{pmatrix} \text{Val} = 1.51 \\ \text{Precision} = \emptyset \\ \text{Unit} = \text{€} \\ \text{Focus} = \text{1L of gas} \\ \text{Date} = \textbf{September 2005} \\ \text{Place} = \textbf{Paris} \\ \text{Restriction} = \emptyset \\ \text{Variation} = \emptyset \end{pmatrix} \quad a_2 \begin{pmatrix} \text{Val} = 1.29 \\ \text{Precision} = \emptyset \\ \text{Unit} = \text{€} \\ \text{Focus} = \text{1L of gas} \\ \text{Date} = \emptyset \\ \text{Place} = \textbf{Toulouse} \\ \text{Restriction} = \emptyset \\ \text{Variation} = \emptyset \end{pmatrix}$$

Figure 4: Example of variation

Variation criteria of numerical values are learnt from texts but can also be infered (or confirmed), for some domains, by common knowledge. For example, the triangle inequality states that for any three points x, y, z, we have:

distance(x, y) $\leq$ distance(x, z) + distance(z, y). From this, we can infer that a distance between two points vary according to restriction (itinerary).

In the following sections, we focus on numerical values which vary according to time.

3.3 Variation mode

The last step consists in identifying the variation mode of values. The idea is to draw a trend (increase, decrease, ...) of variaton in time so that an explanation can be generated. For this purpose, we have a set of couples (numerical value, date) representing the set of extracted answers. From this set, we draw a regression line (a line which comes as close to the points as possible) which determines the relationship between the two variables *value* and *date*.

The **correlation** between two variables reflects the degree to which the variables are related. In particular, Pearson's correlation (r) reflects the degree of linear relationship between two variables. It ranges from $+1$ to -1. A correlation of $+1$ means that there is a perfect positive linear relationship between variables. For example, figure 5 shows that a positive Pearson's correlation implies a general increase of values (trend) whereas a negative Pearson's correlation implies a general decrease. On the contrary, if r is low ($-0.6 < r < 0.6$), then the trend (increase or decrease) is mathematically considered as random[2].
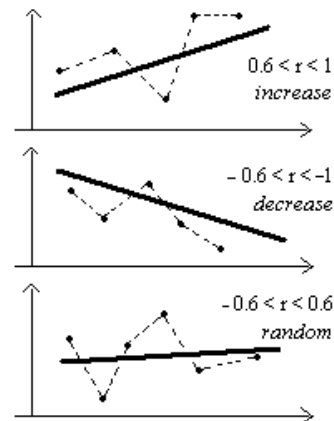

Figure 5: Variation mode

This method determines the variation mode of numerical values (it gives a variation trend) and determines if the values are strongly dependent on time or not (the highest r is, the more the numerical values are dependent on time).

Figure 6 shows the results for the question *How many inhabitants are there in France?* Different numerical values and associated dates are extracted from web pages. The Pearson's correlation is 0.682 meaning that the number of inhabitants increases according to time (between 1999 and 2005).

4 Answer generation

Once the searched numerical values have been extracted and characterized by their variation criteria and mode, a cooperative answer is generated in natural language. It is composed of two parts:
   1. a direct answer if available,
   2. an explanation of the value variation.

---

[2]Statistical Methods for Research Workers, R. Fisher (1925)

```
1999: 60082000    2004: 61000000
2000: 60186184          61700000
2001: 61000000          62000000
2003: 59695000    2005: 62000000
```
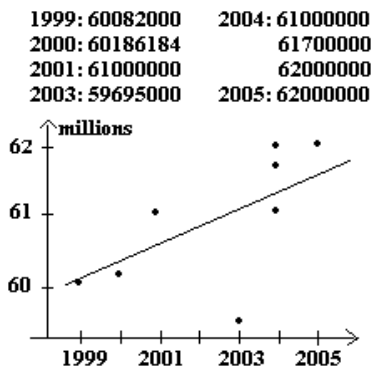
Figure 6: Variation mode: *How many inhabitants are there in France?*

In the following sections, we present some prerequisites to the construction of each of these parts in term of resources and knowledge.

4.1 Direct answer generation

There are mainly two cases: either one or several criteria are constrained by the question (as in *How many inhabitants are there in France in 2005?* where criteria of place and time are given), or some criteria are omitted (or implicit) (as in *How many inhabitants are there in France?* where there is no information on time). In the first case, the numerical value satisfying the constraints is chosen (unification between the criteria of the question and those extracted from web pages). In the second case, we assume that the user wants to have the most recent information.

We focus here on answers which vary according to time. Aberrant values are first filtered out by applying classical statistical methods. Then, when there is only one numerical value which satisfies the temporal constraint (given by the question or the most recent date), then the direct answer is generated from this value. When there is no numerical value satisfying the temporal constraint, only the second part of the answer (explanation) is generated.
In the case of several numerical values satisfying the temporal constraint, there may be approximate values. For example, the following answers (cf figure 6) are extracted for the question *How many inhabitants were there in France in 2004?*:
(1) *61.7 millions: number of inhabitants in France in 2004*.
(2) *In 2004, the French population is estimated to 61 millions*.

(3) *There are 62 millions of inhabitants in France in 2004.*

Each of these values is more or less approximate. The goal is then to identify which values are approximate and to decide which numerical value can be used for the generation task.

For that purpose, we proposed to 20 subjects a set of question-answer pairs. For each question, subjects were asked to choose one answer among a set of precise and approximate values and to explain why. For the previous question, 75% of the subjects prefer answer (1) because it is the most precise one, even if they consider it as an approximation. In majority, subjects explain that an approximate value is sufficient for great numbers and that values must not be rounded up (they proposed *61.7 millions* or *almost 62 millions* as an answer). On the contrary, subjects do not accept approximate values in the financial domain (price, salary, ...) but rather prefer an interval of values.

Thus, the direct answer is generated from the most precise numerical value if available. If all values are approximate, then the generated answer has to explain it: we plan to use prepositions of approximation (*about, almost, ...*) or linguistics clues which have been extracted from web pages (*precision* in the frames). The choice of a particular preposition depends on the degree of precision/approximation of numerical values: PrepNet (Saint-Dizier, 2005) provides a relatively deep description of preposition syntactic and semantic behaviours.

4.2 Explanation generation

Obviously, the generation of the cooperative part of the answer is the most complex because it requires complex lexical knowledge. We present briefly some of the necessary lexical resources. For example, verbs can be used in the answer to express numerical variations. Lexical descriptions are necessary and we use for that purpose a classification of French verbs (Saint-Dizier, 1999) based on the main classes defined by WordNet. The classes we are interested in for our task are mainly those of verbs of change (*increase, decrease*, etc.: in total, 262 verbs in French) and of verbs of movement (*climb, move forward/backward*, etc.: in total, 252 verbs in French) used metaphori-

cally (Moriceau and Saint-Dizier, 2003). From these classes, we have characterized sub-classes of growth, decrease, etc., so that the lexicalisation task is constrained by the type of verbs which has to be used according to the variation mode (if verbs are extracted from web pages as linguistics clues of *variation*, they can also be reused in the answer).

A deep semantics of verbs (change, movement) is necessary to generate an answer which takes into account the characteristics of numerical variation as well as possible: for example, the variation mode but also the speed and range of the variation. Thus, for each sub-class of verbs and its associated variation mode, we need a refined description of ontological domains and selectional restrictions so that an appropriate verb lexicalisation can be chosen: which verb can be applied to prices, to age, etc.? We plan to use proportional series representing verb sub-classes according to the speed and amplitude of variation. For example, the use of *climb* (resp. *drop*) indicates a faster growth (resp. decrease) than *go up* (resp. *go down*): the verb *climb* is prefered for the generation of *The increase of gas prices climb to 20.3% in october 2005* whereas *go up* is prefered in *The increase of gas prices go up to 7.2% in september 2005*.

As for direct answer generation, verbs can possibly be associated with a preposition that refines the information (*The average age of marriage increased by* **about** *5.5 years between 1972 and 2005*).

5 Conclusion

In this paper, we presented an approach for the generation of cooperative numerical answers in a question-answering system. Our method allows us to generate:

(1) a correct synthetic answer over a whole set of data and,

(2) a cooperative part which explains the variation phenomenon to the user,

whenever several numerical values are extracted as possible answers to a question. Information is first extracted from web pages so that numerical values can be characterized: variation criteria and mode are then identified in order to generate explanation to the user. Several future directions are obviously considered:

- an analysis of needs for common knowledge so that the answer characterization task is made easier,

- an analysis of how restrictions are lexicalized in texts (adjectives, relative clauses, etc.) in order to extract them easily,

- an evaluation of the knowledge costs and of what domain specific is (especially for common knowledge about restrictions),

- an evaluation of the quality of answers proposed to users and of the utility of a user model for the selection of the best answer.

References

V. Dahl and H. Abramson. 1984. *On Gapping Grammars*. Proceedings of the Second Logic Programming Conference.

S. Harabagiu and F. Lacatusu. 2004. *Strategies for Advanced Question Answering*. Proceedings of the Workshop on Pragmatics of Question Answering at HLT-NAACL 2004.

D. Maurel and O. Piton. 1999. *Un dictionnaire de noms propres pour Intex: les noms propres géographiques*. Lingvisticae Investigationes, XXII, pp. 277-287, John Benjamins B. V., Amsterdam.

D.L. McGuinness and P. Pinheiro da Silva. 2004. *Trusting Answers on the Web*. New Directions in Question-Answering, chapter 22, Mark T. Maybury (ed), AAAI/MIT Press.

D. Moldovan, C. Clark, S. Harabagiu and S. Maiorano. 2003. *COGEX: A Logic Prover for Question Answering*. Proceedings of HLT-NAACL 2003.

V. Moriceau and P. Saint-Dizier. 2003. *A Conceptual Treatment of Metaphors for NLP*. Proceedings of ICON.

V. Moriceau. 2005. *Answer Generation with Temporal Data Integration*. Proceedings of ENLG'05.

D.R. Radev and K.R. McKeown. 1998. *Generating Natural Language Summaries from Multiple On-Line Sources*. Computational Linguistics, vol. 24, issue 3 - Natural Language Generation, pp. 469 - 500.

P. Saint-Dizier. 1999. *Alternations and Verb Semantic Classes for French*. Predicative Forms for NL and LKB, Kluwer Academic.

P. Saint-Dizier. 2005. *PrepNet: a Framework for Describing Prepositions: preliminary investigation results*. Proccedings of IWCS'05.

B. Webber, C. Gardent and J. Bos. 2002. *Position statement: Inference in Question Answering*. Proceedings of LREC.