

An Analysis of Multi-Focus Questions

Chuan-Jie Lin

Department of Computer Science and Engineering
National Taiwan Ocean University
No 2, Pei-Ning Road, Keelung 202, Taiwan R.O.C.
+886-2-24622192 ext. 6610

cjlin@ntou.edu.tw

Ren-Rui Liu

Department of Computer Science and Engineering
National Taiwan Ocean University
No 2, Pei-Ning Road, Keelung 202, Taiwan R.O.C.
+886-2-24622192 ext. 6610

freesia@cyber.cs.ntou.edu.tw

ABSTRACT

This paper proposes a new definition of question classification based on question focus. Multi-focus questions are defined as questions containing multiple foci. It would be better to present their answers in a table. Besides defining the focus degree, relations between question foci are also discussed. A multi-focus question can be decomposed into several subquestions. The characteristics of the subquestions determine the representation of the final answers. Several simulation experiments have been carried out to verify the importance of focus analysis for multi-focus questions in a QA system.

Categories and Subject Descriptors

H.3.4 [Systems and Software]: Question-answering (fact retrieval) systems

General Terms

Theory, Design, Experimentation

Keywords

Multi-focus questions, focus degree, question analysis, question decomposition

1. INTRODUCTION

Question answering has been a hot research topic in recent years. Large scale QA evaluation projects such as TREC QA-Track [1], QA@CLEF [2], and NTCIR QAC and CLQA Tracks [3][4] greatly support the developments of QA techniques.

In recent QA research, three classes of questions, factoid, list, and complex questions, are generally studied. Factoid questions are commonly studied in famous QA evaluation tasks. They are often “fact-based, short-answer” questions as described in the overview of TREC QA Tracks [5]. For a factoid question, the required answer is often a short string, i.e. a name of an entity, a temporal expression, or a noun phrase, *etc.* For example, the answer to the factoid question “*in what year did the Titanic sink*” is a short temporal expression “1912”.

List questions created in recent QA evaluation forums are also factoid questions. The difference is that a list question often expresses the multiplicity of its answers explicitly (e.g. “*what countries have been struck by Tsunamis*”) or directly requests an amount of various answers (such as “*name 20 countries that produce coffee*”). Answers to a list question are often presented in a list. For example, the list “*Papua New Guinea, Indonesia, New Guinea, Japan*” is an answer to the list question “*what countries have been struck by Tsunamis*”.

Complex questions, on the other hand, often refer to long-answer questions. It means that an answer to a complex question is often a longer passage, say, a set of sentences, a paragraph, or even an article. Questions asking about reason (“*why does the moon turn orange*”), definition (“*what is an atom*”), method (“*how to sort data in Excel*”), or biography (“*who was Galileo*”) are complex questions.

But during our studying on question answering, we found that some questions seem not be able to be classified into the above three classes. An example of such questions is:

Q₁: What are the populations of the countries in the world?

Its answer is not simply a number (as an answer to the factoid question “*what is the population of Japan*”) nor a list of country names (as the answers to the list question “*what are the countries in the world*”), but rather a set of <country_name, population> tuples. It is also suitable to present the answers by a table. For question Q₁, we can give answers in a two-column table like Table 1 whose columns represent “country name” and “population”, respectively.

Table 1. The answers¹ to Q₁
“*What are the populations of the countries in the world?*”

Country	Population
Brazil	178,470
Czech	10,236
Egypt	71,931
Iceland	290
Malaysia	24,425
Yemen	20,010
...	...

A main difference between the new question class and the known three question classes is that: a question of the new class has two or more question foci, while a question of the known classes often has only one question focus. To our best knowledge, there is no research discussing on this topic.

Question decomposition is important for a QA system to answer multi-focus questions. However, the previous research on question decomposition rarely discussed the multiplicity of question foci. Saquete et al. [6] aimed on complex temporal questions expecting only single answers (single-focus questions

¹ Data in Table 1 come from the following UN webpage: http://www.un.org/Pubs/CyberSchoolBus/infonation/e_infonati_on.htm.

as defined later in this paper). The purpose of question decomposition in their work was to resolve temporal information carried in a question. Harabagiu et al. [7] worked on complex questions in LREC which were all single-focus questions. Their decomposition process was more likely a procedure to produce lots of queries which were semantically related to the original question in order to gather more information of answers. Besides, the responses of their system were summaries from multiple documents, not a set of tuples. Lin and Cho [8] proposed a method to do question segmentation, but the main purpose of their work was to identify different questions occurring in a post, not inside one question sentence.

The question decomposition in the START system [9] is most relevant to this topic. If START cannot find answers to a question, it will try to decompose the question and solve the subquestions one by one. For an example selected from their paper, in order to answer the question “*when was the 20th president of the U.S. born*”, START first seeks the answer to the first subquestion “*who was the 20th president of the U.S.*”, whose answer is “*James Abram Garfield*”, and then finds the answer to the second subquestion “*when was James Abram Garfield born*”, which is also the final answer, “*Nov. 19, 1831*”. We can see that the answer to the original question is still a single string. START does not yet manage to answer in table fashion.

Katz et al. [9] referred to a question needed to be decomposed as a “complex question”. In our definition, such a question is classified as a multi-focus single-response question.

This paper proposes a new principle to classify questions which have more than one question focus. Section 2 introduces the idea of multi-focus questions by defining the focus degree and the relations among the foci. Section 3 gives some simulation experiments to show the importance of handling multi-focus questions. Section 4 concludes this paper.

2. Definition of Multi-Focus Questions

The definitions of question focus are not quite the same in many papers. Lehnert [10] defined it as “*the question concept that embodies the information expectations expressed by the question*”, while recent QA research groups [11][12] often refer to it as “*a word or sequence of words which indicate what information is being asked for in the question*” which means a question focus is a substring of a question.

To extend the definition to multi-focus questions, we first consider the necessity of question decomposition, and then define the question focus degree.

2.1 Question Decomposition

As we have seen in Section 1, some questions need to be decomposed before looking for answers. The decomposition process divides a question into a set of subquestions.

In our observation, a subquestion may depend on another question, i.e. a subquestion cannot be answered before another subquestion is answered. The dependencies of the subquestions form a dependency chain. Take Q_1 as an example:

Q_1 : What are the populations of the countries in the world?

$Q'_{1,1}$: What are the countries in the world?

$Q'_{1,2}$: What is the population of <ans $Q'_{1,1}$ >?

Dependency chain: $Q'_{1,1} \rightarrow Q'_{1,2}$

Question Q_1 is decomposed into two subquestions and its second subquestion depends on its first subquestion, which means we have to find the names of all the countries in advance so that we can answer the population of each country.

In the next sections of this paper, we call the rightmost subquestion in a dependency chain the *final subquestion*, while other subquestions are *preceding subquestions*. Note that there can be more than one final subquestion in a question, as we will see later.

2.2 Question Focus and Focus Degree

Our definition of question focus of a subquestion follows the convention. The **question focus of a question** can be re-defined as the set of the question foci of its subquestions.

Now we define the **focus degree** (FD) as the cardinality of the question focus set. For example:

Q_2 : Give me the capitals, national flowers, and national trees of all the countries.

The question foci of Q_2 are “*capital*”, “*national flower*”, “*national tree*”, and “*country*”. Therefore, the focus degree of question Q_2 is 4.

A **multi-focus question** is a question whose focus degree is larger than 1. On the other hand, a question with a focus degree as 1 is called a **single-focus question**.

2.3 Relations between Question Foci

In order to learn the characteristics of multi-focus questions, we visited three QA websites, PTT², Yahoo Knowledge⁺³, and Baidu Zhidao⁴, to collect multi-focus questions. Although the questions are written in Chinese, we believe that the nature of multi-focus questions should be language-independent, so the Chinese question set can reveal the characteristics as the questions written in other languages.

Because the question collections in those QA websites are too large to browse, we submitted some particular queries which might often be realized by multi-focus questions. Examples of the queries are “*country + caption*” and “*city + population*”.

After browsing 2,650 questions, we collected a set of 222 multi-focus questions. The number does not reveal the true ratio of multi-focus questions in human languages. A larger scale of investigation should be conducted in the future.

In our investigation, we find that the relation between two question foci is helpful to define the format of required answers. Note that when discussing the relation between two question foci, only their head nouns or the main constituents are considered.

The following four relations are explored from the collected multi-focus question set:

² <http://www.ptt.cc/>

³ <http://tw.knowledge.yahoo.com/>

⁴ <http://zhidao.baidu.com/>

(1) Coordinate Relation

The original question is a combination (mostly coordinate conjunction) of its subquestions. For example:

Q₃: I want to know the biggest city and the largest port in Japan.

Q'_{3,1}: What is the biggest city in Japan?

QF₁

Q'_{3,2}: What is the largest port in Japan?

QF₂

Dependency chain: {Q'_{3,1}, Q'_{3,2}}

Q₃ is in fact a conjunction of its two subquestions, Q'_{3,1} and Q'_{3,2}. In this case, two subquestions can be answered independently.

(2) Entity-Attribute Relation

One question focus represents an entity type while the other focus represents an attribute of the entity type. For example:

Q₄: What are the 5 largest cities and their populations?

Q'_{4,1}: What are the 5 largest cities?

QF₁

Q'_{4,2}: What is the population of <ansQ'_{4,1}>?

QF₂

Dependency chain: Q'_{4,1} → Q'_{4,2}

QF₂ “*population*” in Q₄ is an attribute of QF₁ “*city*”. To answer this question, we should know the 5 largest cities before we can find the population of each city.

(3) Entity-Relationship Relation

Both question foci are entity types. One focus is explicitly expressed but the other focus is hidden. Instead, the surface text of the other focus represents the relationship between the two entity types. For example,

Q₅: Names of capitals of Spanish speaking countries?

Q'_{5,1}: What are Spanish speaking countries?

QF₁

Q'_{5,2}: Name of capital of <ansQ'_{5,1}>?

QF₂

Dependency chain: Q'_{5,1} → Q'_{5,2}

The two question foci in Q₅ are actually “*city*” and “*country*”. The phrase QF₂ is the relationship “*capital-of*” between the explicitly stated focus QF₁ “*country*” and the hidden focus “*city*”.

Someone may argue that a question like Q₅ does not need a table to present its answers, because the question only requests a list of capital names. But it seems to us that a list of <*country*, *capital*> tuples is more expressive and welcome.

Because there are also relationships relating to three or more entities, naturally there will be questions asking all the entities involved in a specific relationship at the same time. This definition can be extended to a relation among three or more entities.

(4) Thematic Relation

A more complicated case is that the two question foci are involved in a predicate-argument structure. “Argument” here refers to semantic argument. The two foci are often two arguments related to a predicate. For example:

Q₆: When are the proper times to take vitamins?

Q'_{6,1}: List all the vitamins.

QF₁

Q'_{6,2}: When is the proper time to take <ansQ'_{6,1}>?

QF₂

Dependency chain: Q'_{6,1} → Q'_{6,2}

This question asks for a list of vitamins and the best time to take each kind of vitamin. Take the predicate “take” into account, QF₁ “*vitamin*” is the patient of the predicate and QF₂ “*proper time*” is the time of the predicate. Both are the semantic arguments of the predicate “take”.

Here is another example:

Q₇: How to take care of fruit trees?

Q'_{7,1}: List all the fruit trees.

QF₁

Q'_{7,2}: How to take care of <ansQ'_{7,1}>?

QF₂: *the procedure to take care*

Dependency chain: Q'_{7,1} → Q'_{7,2}

The surface text of QF₂ is “*how to take care*” which asks for a procedure to take care of something. Considering the action of “*taking care*”, QF₂ corresponds to its method and QF₁ “*fruit trees*” is the beneficiary. Different fruit trees may need different care procedures, so the answers could be presented by multiple columns.

This definition can also be extended to a relation among two or more entities, because there can be many arguments related to a predicate.

It also seems possible to define a relation that one focus is a predicate and the other focus is one of its arguments, just like in the question “who has done what to whom”. But we cannot find a good example to illustrate this case, or perhaps it is because people rarely ask questions without predicates.

According to the definitions given above, relations of all pairs of question foci in a question can be determined. Take Q₂ as an example:

Q₂: Give me the capitals, national flowers, and national trees of all the countries.

Q'_{2,1}: List all the countries.

QF₁

Q'_{2,2}: What is the capital of <ansQ'_{2,1}>?

QF₂

Q'_{2,3}: What is the national flower of <ansQ'_{2,1}>?

QF₃

Q'_{2,4}: What is the national tree of <ansQ'_{2,1}>?

QF₄

Dependency chain: Q'_{2,1} → {Q'_{2,2}, Q'_{2,3}, Q'_{2,4}}

The relations between QF₁/QF₂, QF₁/QF₃, and QF₁/QF₄ are all entity-relationship relations (relationships between “*country*” and

“city”, “flower”, and “tree”, respectively). The relations between QF_2/QF_3 , QF_2/QF_4 , and QF_3/QF_4 are coordinate relations.

Note that not all pairs of foci in a question have relations defined above. For example:

- Q₈: List the capitals of the countries in the world and their populations.
 Q'_{8,1}: List all the countries.
 QF_1
 Q'_{8,2}: What is the capital of <ansQ'_{8,1}>?
 QF_2
 Q'_{8,3}: What is the population of <ansQ'_{8,2}>?
 QF_3
 Dependency chain: Q'_{8,1} → Q'_{8,2} → Q'_{8,3}

The foci QF_1 “country” and QF_2 “capital” have entity-relationship relation. QF_2 “capital” and QF_3 “population” have entity-attribute relation. But QF_1 and QF_3 do not have any direct relation.

2.4 Answer Format

The new definition of question classes is neither incompatible nor exclusive to the conventional definition. In fact, they define questions in different aspects.

A multi-focus question is defined by the number of question foci. A factoid or complex question is defined by the complexity of its answer types. A list question is defined by the multiplicity of its requested answers. We call a question a **single-response question** if it does not request a list of answers.

If a question can be decomposed into several subquestions, the

Q: What are the names of all the countries in Europe along with their capitals?
 A: I don't know the answer.
 Q: What are all the countries in Europe?
 A: I have information about the following countries in Europe:
 Albania, Andorra, Austria, Belarus, Belgium, Bosnia and Herzegovina, the UK, ..., Switzerland, and Ukraine
 Q: What is the capital of Albania?
 A: Tirane
 Q: What is the capital of Andorra?
 A: Andorra La Vella
 ...
 Q: What is the capital of Ukraine?
 A: Kiev

Final Answers:

Country	Capital
Albania	Tirane
Andorra	Andorra La Vella
Austria	Vienna
Belarus	Minsk
Belgium	Brussels
...	...
Switzerland	Bern
Ukraine	Kiev

Figure 1. Answer Responses to TQ₁ “What are the names of all the countries in Europe along with their capitals?”

complexity and multiplicity of each subquestion’s answer type determines the format of the answers.

Considering a question Q which can be decomposed into subquestions $Q_1, Q_2 \dots Q_n$, and the dependency chain is $Q_1 \rightarrow Q_2 \rightarrow \dots \rightarrow Q_n$. The answers to the question Q can be presented in an n -column table where the i th column gives the answers to the subquestion Q_i and the data in this column match the answer type of Q_i .

If Q_1 is a list question, the number of possible answers shown in the first column will be larger than 1, so the table will contain multiple rows. But if all the preceding subquestions Q_i (where $1 \leq i \leq n-1$) are single-response questions, there is only one possible combination of data in the first $n-1$ columns, so the original question Q becomes a conventional single-response question or a list question according to the type of the final subquestion Q_n .

3. Simulation Experiments

In order to show the necessity of handling multi-focus questions, three simulation experiments have been carried out and the results confirm our assertion.

Because the structures (in syntax level or semantic level) of multi-focus questions have not been fully studied, we have no modules to decompose multi-focus questions yet. In order to see the effect of multi-focus handling in a QA system, we manually performed the question decomposition and submitted the subquestions to an online QA system. We chose the START system [9] to do the experiment.

Because START is good at answering geographical questions, three multi-focus questions asking information of countries are selected as the testing questions. They are:

- TQ₁: What are the names of all the countries in Europe along with their capitals?
 TQ₂: What are the climates of the countries in North America?
 TQ₃: Can anyone give me a list of the largest 8 countries in Asia with their flags?

All their first subquestions are factoid list questions. All of the final subquestions are single-response questions, but their answer types are different. The final subquestion of TQ₁ is a factoid question, while the one of TQ₂ is more like a complex question. The final subquestion of TQ₃ can be considered as a factoid question only that its answer is a picture. The simulation experiments are described in the following subsections.

3.1 Simulation Experiment 1

Figure 1 illustrates the procedure of answering the multi-focus question TQ₁ by using START. Question TQ₁ was manually decomposed as follows:

- TQ₁: What are the names of all the countries in Europe along with their capitals?
 TQ'_{1,1}: What are all the countries in Europe?
 TQ'_{1,2}: What is the capital of <ansTQ'_{1,1}>?

Answers to the subquestion TQ'_{1,1} were successfully collected from START system. Each answer was then used to replace <ansTQ'_{1,1}> to create its corresponding subquestion TQ'_{1,2}. For example, “Albania” was one of the answers to the first

subquestion. It created a second subquestion “*what is the capital of Albania*” and then submitted to START. The answer to the second subquestion was paired with the answer to the first subquestion and formed a tuple in the table of final answers.

As we can see, when the original question TQ₁ is submitted to START without decomposition, the response of START is “*I don't know the answer.*” By detecting question foci and decomposing the question properly, answers to TQ₁ can be wonderfully created.

3.2 Simulation Experiment 2

Question TQ₂ was manually decomposed as follows:

- TQ₂: What are the climates of the countries in North America?
- TQ'_{2,1}: What are the countries in North America?
- TQ'_{2,2}: What is the climate of <ansTQ'_{2,1}>?

Q: What are the climates of the countries in North America?																		
A: Unfortunately, I don't know what the climates of countries are in North America.																		
Q: What are the countries in North America?																		
A: I have information about the following countries in North America: the Bahamas, Barbados, Bermuda, Canada, Clipperton Island, Cuba, the Dominican Republic, Greenland, Grenada, Haiti, Jamaica, Mexico, Puerto Rico, Saint Pierre and Miquelon, and the USA																		
Q: What is the climate of the Bahamas?																		
A: tropical marine; moderated by warm waters of Gulf Stream																		
Q: What is the climate of Barbados?																		
A: tropical; rainy season (June to October)																		
...																		
Q: What is the climate of the USA?																		
A: mostly temperate, but tropical in Hawaii and Florida, arctic in Alaska, semiarid in the great plains west of the Mississippi River, ...																		
Final Answers:																		
<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="text-align: left;">Country</th> <th style="text-align: left;">Climate</th> </tr> </thead> <tbody> <tr> <td>the Bahamas</td> <td>tropical marine; moderated by warm waters of Gulf Stream</td> </tr> <tr> <td>Barbados</td> <td>tropical; rainy season (June to October)</td> </tr> <tr> <td>Bermuda</td> <td>subtropical; mild, humid; gales, strong winds common in winter</td> </tr> <tr> <td>Canada</td> <td>varies from temperate in south to subarctic and arctic in north</td> </tr> <tr> <td>Clipperton Island</td> <td>tropical; humid, average temperature 20-32 degrees C, wet season (May to October)</td> </tr> <tr> <td>...</td> <td>...</td> </tr> <tr> <td>Saint Pierre and Miquelon</td> <td>cold and wet, with much mist and fog; spring and autumn are windy</td> </tr> <tr> <td>the USA</td> <td>mostly temperate, but tropical in Hawaii and Florida, arctic in Alaska, semiarid in the great plains west of the Mississippi River, ...</td> </tr> </tbody> </table>	Country	Climate	the Bahamas	tropical marine; moderated by warm waters of Gulf Stream	Barbados	tropical; rainy season (June to October)	Bermuda	subtropical; mild, humid; gales, strong winds common in winter	Canada	varies from temperate in south to subarctic and arctic in north	Clipperton Island	tropical; humid, average temperature 20-32 degrees C, wet season (May to October)	Saint Pierre and Miquelon	cold and wet, with much mist and fog; spring and autumn are windy	the USA	mostly temperate, but tropical in Hawaii and Florida, arctic in Alaska, semiarid in the great plains west of the Mississippi River, ...
Country	Climate																	
the Bahamas	tropical marine; moderated by warm waters of Gulf Stream																	
Barbados	tropical; rainy season (June to October)																	
Bermuda	subtropical; mild, humid; gales, strong winds common in winter																	
Canada	varies from temperate in south to subarctic and arctic in north																	
Clipperton Island	tropical; humid, average temperature 20-32 degrees C, wet season (May to October)																	
...	...																	
Saint Pierre and Miquelon	cold and wet, with much mist and fog; spring and autumn are windy																	
the USA	mostly temperate, but tropical in Hawaii and Florida, arctic in Alaska, semiarid in the great plains west of the Mississippi River, ...																	

Figure 2. Answer Responses to TQ₂, “*What are the climates of the countries in North America?*”

Figure 2 illustrates the procedure of answering the multi-focus question TQ₂ by using START. Again, START failed to answer the original multi-focus question, but then successfully provided the final answers following our answering procedure.

3.3 Simulation Experiment 3

TQ₃ is a more demonstrative example, as its answers are pictures, not text. It was manually decomposed as follows:

- TQ₃: Can anyone give me a list of the largest 8 countries in Asia with their flags?
- TQ'_{3,1}: What are the largest 8 countries in Asia?
- TQ'_{3,2}: What is the flag of <ansTQ'_{3,1}>?

Figure 3 illustrates the procedure of answering the multi-focus TQ₃ by using START. After submitting TQ₃, START tried hard to find information to resolve the anaphora but failed in the end. Our proposed procedure successfully created the final answers once more.

4. Conclusion and Future Works

This paper proposes a new analyzing method to define multi-focus questions. The focus degree is used to describe the number of question foci of a question. Four relations between question foci are defined and explained. They can be used to analyze a multi-focus question.

Three simulation experiments were conducted to verify the necessity of handling multi-focus questions. Without multi-focus question processing, none of the testing questions could be answered by an online QA system. But if the relations between the answers of the subquestions were correctly identified and the subquestions were correctly created, wonderful responses could be constructed.

The most important skills to be developed in the future are relation identification and question decomposition for multi-focus questions. A large set of multi-focus questions is needed to learn the structures or patterns in multi-focus questions. A real QA system designed to answer multi-focus questions is needed to be developed to see the actual effect.


Furthermore, whether a question contains multiple foci is also ambiguous. Take question Q₆ as an example. It is possible that Q₆ requests different answers for different vitamins, or simply asks for one proper time for all kinds of vitamins. It remains a future work to find a method to solve the ambiguity.


5. References

- [1] Dang, H. T., Lin, J., and Kelly, D. 2006. Overview of the TREC question answering track. In Proceedings of the 15th Text REtrieval Conference (TREC 2006).
- [2] Giampiccolo, D., Forner, P., Peñas, A., Ayache, C., Cristea, Jijkoun, D. V., Osenova, P., Rocha, P., Sacaleanu, B., and Sutcliffe R. 2007. Overview of the CLEF 2007 multilingual question answering track. In Working Notes for the CLEF 2007 Workshop (Budapest, Hungary, September 19-21, 2007).

Q: Can anyone give me a list of the largest 8 countries in Asia with their flags?
 A: I was unable to find the referent for the pronoun THEIR. Sorry.
 Q: What are the largest 8 countries in Asia?
 A: The following are the countries ranked in the top eight places for area among countries in Asia.

- China
- India
- Kazakhstan
- Saudi Arabia
- Indonesia
- Iran
- Mongolia
- Pakistan

Q: What is the flag of China?
 A: 

Q: What is the flag of India?
 A: 

...
 Q: What is the flag of Pakistan?
 A: 

Final Answers:







Country	Flag
China	
India	
Kazakhstan	
Saudi Arabia	
Indonesia	
Iran	
Mongolia	
Pakistan	

Figure 3. Answer Responses to TQ₃, “Can anyone give me a list of the largest 8 countries in Asia with their flags?”

[3] Fukumoto, J., Kato, T., Masui, F., and Mori, T. 2007. An overview of the 4th question answering challenge (QAC-4) at NTCIR Workshop 6. In Proceedings of the 6th NTCIR Workshop Meeting on Evaluation of Information Access Technologies: Information Retrieval, Question Answering and Cross-Lingual Information Access (Tokyo, Japan, May 15-18, 2007).

[4] Sasaki, Y., Lin, C-J., Chen, K-H., and Chen, H-H 2007. Overview of the NTCIR-6 cross-lingual question answering (CLQA) task. In Proceedings of the 6th NTCIR Workshop Meeting on Evaluation of Information Access Technologies: Information Retrieval, Question Answering and Cross-Lingual Information Access (Tokyo, Japan, May 15-18, 2007).

[5] Voorhees, E. 2001. Overview of the TREC 2001 question answering track. In Proceedings of the Tenth Text REtrieval Conference (TREC 2001).

[6] Saquete, E., Martinez-Barco, P., Munoz, R., and Gonzalez, J. L. V. 2004. Splitting complex temporal questions for question answering systems. In Proceedings of ACL 2004, 566-573.

[7] Harabagiu, S., Lacatusu, F., and Hickl, A. 2006. Answering complex questions with random walk models. In Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, 220-227.

[8] Lin, C-J. and Cho, C-H. 2006. Question pre-processing in a QA system on Internet discussion groups. In Proceedings of COLING-ACL 2006, Workshop on Task-Focused Summarization and Question Answering, 16-23.

[9] Katz, B., Borchardt, G., and Felshin, S. 2005. Syntactic and semantic decomposition strategies for question answering from multiple resources. In Proceedings of the AAAI 2005 Workshop on Inference for Textual Question Answering, 35-41.

[10] Lehnert, W. 1978. The Process of Question Answering - A Computer Simulation of Cognition, Lawrence Erlbaum Associates, Hillsdale, N.J.

- [11] Moldovan, D., Harabagiu, S., Pasca, M., Mihalcea, R., Girju, R., Goodrum, R., Rus, V. 2000. The structure and performance of an open-domain question answering system. In Proceedings of the Conference of the Association for Computational Linguistics (ACL-2000), 563-570.
- [12] Ferret, O., Grau, B., Hurault-Plantet, M., Illouz, G., Monceaux, L., Robba, I., and Vilnat, A. 2001. Finding an answer based on the recognition of the question focus. In Proceedings of TREC-10, 362-370.