The Influence of Semantic Link Network on the Ability of Question-Answering System

Bei Xu and Hai Zhuge*

Nanjing University of Posts and Telecommunications, Nanjing, China
Aston University, Birmingham, UK

Abstract

Semantic Link Network plays an important role in representing and understanding text. This paper investigates the influence of semantic links on the basic abilities of a type of QA system that extracts answers from a range of texts (answer range). Research concerns how semantic links influence the answer range and the performance of this type of QA system. Research also concerns the ability to answering different types of questions and supporting different patterns of answering questions. Based on the semantic link network extracted from Wikipedia, an experimental QA system is developed to answer questions according to a range of pages in Wikipedia. Research reached the following results: (1) the answer range and the semantic link network influence each other: keeping a certain range of performance, increase one can decrease the request of the other; and, (2) the semantic link network can enhance the ability of QA system in answering questions and supporting patterns of answering questions covered by semantic link network.

1. Introduction

Question answering (QA) system is an advanced information service that inputs question in form of natural language and then outputs answer in response to the question in the same language.

One type of QA system extracts answers from a collection of texts. The performance of this type of QA system is mainly influenced by the following two factors:

The first is the scale of answer range from which a QA system extracts answers. Intuitively, a bigger range of texts provides more potential answers. Some QA systems set answer range on a certain domain to provide specialized question answering [16]. And some QA systems manage to provide domain-free question answering by collecting a large number of texts, such as Watson (https://www.ibm.com/watson/, integrating corpus-based language processing, QA experience, search the evidence to support answer, and machine learning) and Webclopedia [18].

The second is the extent of understanding text. A deeper understanding enables this type of QA system to extract more answers. One indicator of deeper understanding is knowing more semantic objects and semantic

*Corresponding author: h.zhuge@aston.ac.uk

links within text. Some external knowledge base can help identify semantic objects and establish semantic link within texts [24, 28]. Semantic link network is a self-organized semantic model that can describe and store semantic objects and semantic links within texts. It can semantically link various resources and carry out reasoning on semantic links [44, 45, 46, 55]. Semantic Link Network can support different abilities of QA system with the types of semantic links. For example, the semantic link on category can help a QA system answer questions on definition because a definition often uses super-category or subcategory of a semantic object; the semantic link on cause-effect relation can help answer questions on reasoning [52]; the semantic link on position or direction can help answer *where* questions; the semantic link on analogy between objects can help answer questions on metaphor; semantic links on time, space, or process support answering details about *reason* or *procedure* questions.

Previous works seldom analyze the influences of the scales of answer range and types of semantic links on the performance of QA system.

One type of semantic link is the semantic link on category which reflects a way to organize and manage objects. There are three types of semantic link on category: *subcategory* link, *is-part-of* link and *instance* link, which influence QA system on the following aspects.

1. They can influence the range of texts from which a QA system searches answers. A QA system that holds more categories and instances is likely to extract more answers from a certain range of text. But a QA system can only process a small range of texts if it "knows" only a small number of categories and instances.

2. Some types of questions can be answered by these links. For example, they can help answer the questions about concept or specialized/generalized concept as people tend to build a concept through its subcategories, super-categories, instances, parts and whole.

3. They can support a part of the patterns of answering questions. A pattern is the generalization of a set of answers to a question. For a question, different answers may be obtained according to different patterns. For example, for a question "*what is data mining*?", the answer "*data mining is a research field of computer science.*" is obtained by syntactically replacing the interrogative "*what*" with a certain semantic object "*a research field of computer science*", and the answer "*Data mining focuses on discovering patterns in large data sets*" gives the connotation of "*data mining*". The two answers follow different patterns.

Therefore, it is significant to study the influence of the semantic links on the performance of QA system. The influence should be more prominent when a question concerns specialized concepts because specialized concepts usually have rigorous and deep category structure.

This paper builds an experimental QA system based on the Semantic Link Network that consists of the categories and instances extracted from Wikipedia and the three types of semantic link. The system extracts answer from a given range of text. The paper proposes three propositions on the influences of the semantic links and answer range on the performance of QA system and the mutual-influence between them. Research verifies the proposition by adjusting the scale of the semantic link network and the range of text respectively. Research further investigates the abilities of the semantic link network to enhance QA systems in answering different types of question and supporting different patterns of answering question.

2. Related works

2.1. QA system

One type of QA systems generates answers according to the structured or semi-structured knowledge representations about question domains. Some of these QA systems such as Baseball [17] and Lunar [6] use a structured query language to represent query and use database to store domain knowledge to support question answering. These systems require a lot of manual or semi-manual works to acquire domain knowledge [4] and they are fragile at the border of domain knowledge. The form of question has been expanded to natural language as used in Berkeley Unix Consultant [40]. Some systems use question-answer pairs, which can be considered as a kind of semi-structured knowledge, as experience to answer question [13, 41]. Deep learning model has been applied to search answers from big sets of question-answer pairs [7]. Chatbots can be regarded as a kind of QA systems that need to maintain the coherence between questions and answers. Some chatbots extract answers directly from a large number of dialogues which is a kind of semi-structured representation of text. They use rule-based

methods, such as CleverBot [56], or training models, such as RNN-based Encoder-Decoder model [57]. Typical chatbots include *Alicebot* (http://www.alicebot.org/), *Xiaoice* [54], and *Siri* of *Apple*.

The other type of QA systems directly extract answers from free-text. The systems search answers in a large collection of plain texts. Usually, these systems operate with the following steps.

1. *Question analysis*, which classifies questions into several types [27, 36], such as *who*, *how*, *where*, *what*, and *why*. Some QA systems have finer classification of questions. For example, *NLP* laboratory at Southern Methodist University (SMU) classifies questions into 16 types [25]. Typical ways of classification include: Cymfony, ISI, SMU and UIUC (http://trec.nist.gov/). They classify questions according to the types of information required by question. Classifying questions into several well-organized types is proven to be an effective way of improving the performance of QA system [58]. And a QA system's ability can also be evaluated in terms of question types.

2. *Information retrieval*, which locates or narrows the range of candidate answers in a given collection of texts. Frequently used methods include: Boolean model, probability model and Vector Space Model [12, 38]. Some works consider paragraph instead of sentence as the basic unit of answers, such as *MultiText* [8], *SiteQ* [21], and relation-matching method [10].

3. *Answer extraction*, which selects answers from candidate answers. There are four representative methods: (a) using superficial features, like the frequency of answer [25], to calculate the distance between a question and each candidate answer; (b) discovering the relations between a question and candidate answers, and then transforming them into triples or logic representation [11]; (c) using machine learning to train answer patterns [9]; and, (d) calculating the probabilities of an answer's appearance, such as the noisy-channel approach [14] and the probabilistic graphical model [20]. Deep learning has been used in answer selection as well [19].

Free-text based QA systems often require external knowledge base to deal with the semantics of text. One form of the external knowledge representation is ontology that provides a semantic mechanism, usually in triples, for describing the classification relations between semantic objects [29, 30]. The technique is often used in a closed-domain, such as open biomedical ontology [35] and gene ontology [16] for biology domain. Therefore, some QA systems are able to answer specialized questions, such as *EAGLi* (Engine for Question-Answering in Genomics Literature, http://eagl.unige.ch/KART/) for health and life scientists. Another form of the external knowledge is open-domain, such as Wikipedia [26] and Freebase [37].

2.2. Semantic link network

Semantic Link Network (SLN) consists of a set of semantic objects and semantic links between them. Research can trace to the definition of inheritance rules and Active Document Framework [49, 50]. SLN has been applied to various applications such as text summarization, recommendation and semantics-based peer-to-peer search [33, 42, 47], and text summarization [52, 53]. SLN was integrated with a multi-dimensional resource space model to form a combined power for modeling semantics [42, 48, 51]. SLN develops in parallel with Tim Berners-Lee's Linked Data proposed in 2006. It is also a model for implementing the knowledge graph proposed by Google in 2012. Recently, SLN has been used to link things in cyberspace, physical space and social space to form a Cyber-Physical-Social SLN CPSoSLN [55].

Some existing open-domain information bases, such as WordNet and the Wikipedia [26], can help build semantic link network. They are already used in many QA systems.

Some QA systems use WordNet to match common words in question and answer [5], rank and filter answers [31, 23], or expand keywords in question [18]. However, WordNet focuses on describing single word rather than phrases like "*natural language processing*". The resources in Wikipedia contain rich contents contributed by worldwide editors. So, some QA systems use Wikipedia [2, 22, 32] as external knowledge, such as object recognition [20, 26], answer-related sentence identification [1], and query expansion [15]. A multi-layer recurrent neural network model was used to retrieve answers from Wikipedia pages [39]. These QA systems use local information such as the content of pages or the subcategory of given words.

3. Propositions

Answer range and semantic link influence not only the performance of QA system but also each other. Hence we have the following proposition.

Proposition 1 (*Overall influence*). The performance of a QA system can be improved by providing a wider answer range (given more texts that contain potential answers) and more semantic links on the answer range.

In addition to the overall influence, they have the following side influence.

Proposition 2 (*Side influence*). Fixing an answer range and a question, a QA system is more likely to find more answers by providing more semantic links on the answer range. Fixing a semantic link network and a question, more answers can be extracted from a wider answer range with more relevant texts.

To keep a certain performance, answer range and semantic links may influence each other.

Proposition 3 (*Mutual influence*). To keep a certain performance, given a narrower answer range implies that a QA system needs more semantic links on the answer range, and given less semantic links requests a QA system to widen the answer range.

An experimental system based on SLN is built for verifying the above propositions.

4. Constructing Semantic Link Network

4.1. Components

The construction of semantic link network requires the following components.

1. *Answer range*, which is a set of sentences within one text or multiple texts, from which a QA system extracts answers. A semantic link network on text reflects the core semantics of the text(s) so it can cover the main answer range [47][52][53].
2. *Semantic object*, which is semantic image of language unit. Here, it takes the form of a string that indicates a category in Wikipedia. Synonyms are mapped into one semantic object.
3. *Semantic link*, which connects two semantic objects via relations, which are automatically extracted from Wikipedia herein. In Wikipedia, *part-of*, *has-part* or *subclass-of* relations connect pages. One page is subordinate to one or multiple categories; one category contains multiple subcategories and pages; and, one category is subordinate to one or multiple categories. Therefore, here concerns *subcategory*, *is-part-of*, and *instance* semantic links between semantic objects.
4. *Semantic link network*, which consists of semantic objects connected by semantic links. A QA system concerns two kinds of semantic link network: One reflect the global semantics. Herein, it covers the semantic objects in Wikipedia, which is regarded as the knowledge base of our QA system. The other contains the semantic objects within an answer range, which provides local semantics for a QA task. Given a set of semantic objects K, the constructed semantic link network can be represented as $SLN(K)$. The construction process is elaborated in Appendix E.

Fig. 1 illustrates the construction of a part of semantic link network, where "Category: machine learning" and "Page: machine learning" are mapped into one semantic object. The construction follows the rules described in Appendix A for guaranteeing the completeness, unify personalized perspectives, and avoid dead end.
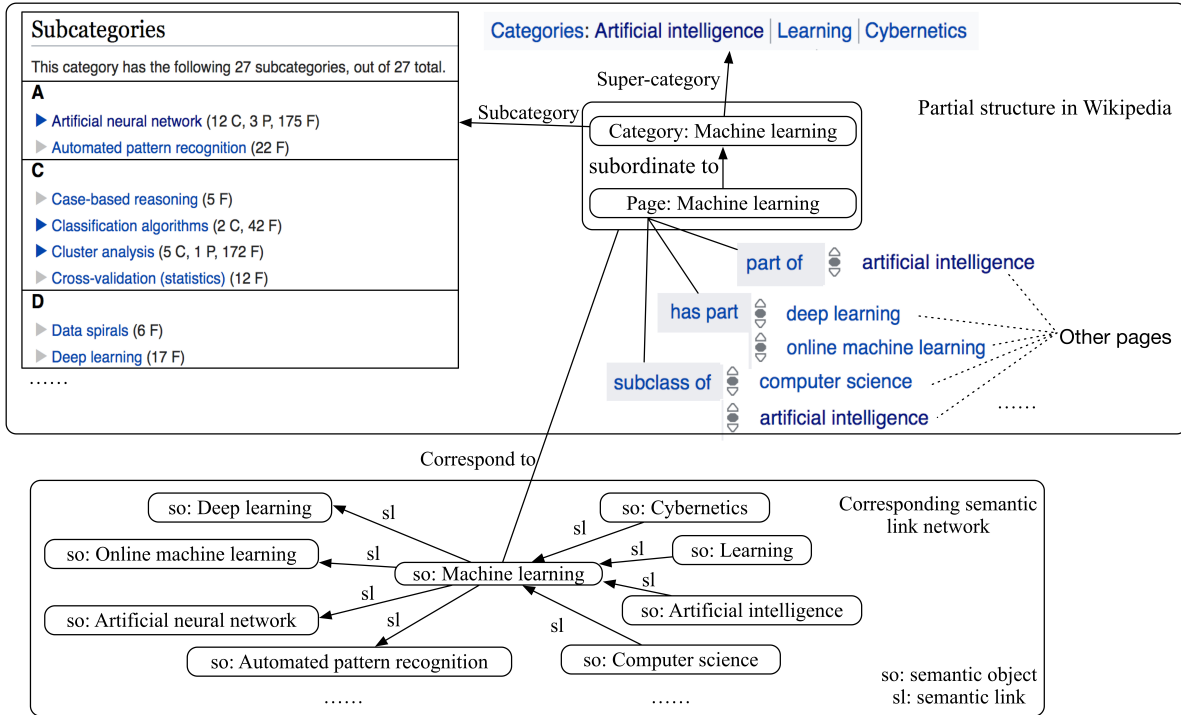
Fig. 1. Construction of semantic link network.

4.2. Function

QA system based on SLN has the following distinguished functions (introduced in detail in Appendix D):
1. *Semantic filtering*, which can narrow answer range by selecting the sentences that matches the type of question.
2. *Expanding questions and answer semantically*, which can provide extended answers by semantically expanding semantic objects in question and answer.
3. *Semantic matching*, which can match answer and question by checking whether the semantic objects in answer and question are connected by semantic links and whether the way to answering the question require the semantic links.

5. Question Answer System

5.1. Architecture

The architecture of the experimental QA system is shown in Fig. 2. It requires two inputs: a question and an answer range. The frames with folded corner represent data. The arrows represent data flow. The rounded rectangle represents a module that processes or provides data. The Question Analyzing Module (QAM) outputs the type of an input question and the semantic objects in the question. It concerns 7 basic types of question: (1) *Object description*, (2) *Target object*, (3) *Time*, (4) *Location*, (5) *Person*, (6) *Reason*, and (7) *Process*, which are elaborated in Appendix B in detail. The Key Terms Extracting Module (KEM) extracts the key terms from an answer range. The Answer Filtering Module (AFM) filters answer range through filtering function. KEM and AFM are elaborated in Appendix F in detail. The Semantic Link Network Construction Model (SCM)

constructs a semantic link network on a given answer range. The Answer Extracting Module (AEM) extracts answers from the filtered answer range.
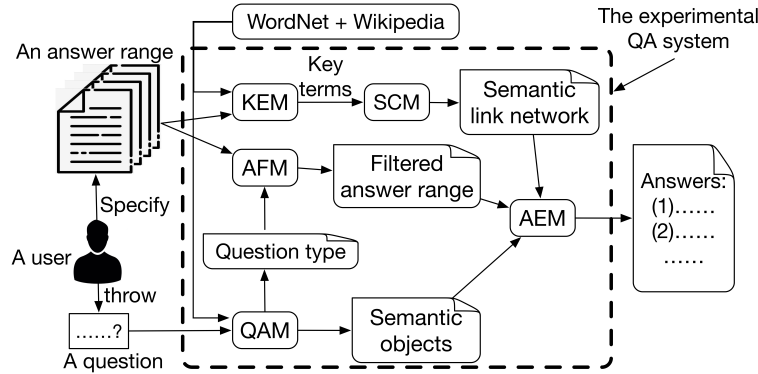

Fig. 2. The architecture of the experimental QA system.

The question analyzing module and the answer filtering module make up a simple QA system. The Answer filtering module, the key terms extraction module, and the semantic link network construction module enhance the system with the semantic link network.

5.2. Answer extracting method

To embody the influence of the semantic link network, the answer to a question is extracted by the introduced three functions, implemented by the answer filtering module and the answer extracting method. The expanding function enables the degree of the relevancy between the extracted answer and the expanded answer to be determined by the length of the semantic link chains between the semantic objects in question and the expanded answer. The matching function enables the degree of matching between question and answer to be determined by the length as well. Therefore, the expanding function and the matching function require the QA system to trace semantic links while extracting answers. So we design *category distance* to reflect the distance between answer and question according to the semantic link network. Given that the semantic objects in a question Q are $oq_1, oq_2 \dots oq_m$, and the semantic objects in a sentence S are $os_1, os_2 \dots os_n$, the *category distance* between Q and S is calculated by equation (1),

$$CD(Q,S) = \frac{1}{mn} \times \sum_{i=1}^{m} \sum_{j=1}^{n} D(oq_i, os_j) \tag{1}$$

where $CD(Q, S)$ represents the *category distance* between Q and S, $CD(oq_i, os_j)$ denotes the *category distance* between two semantic objects oq_i and os_j, which is calculated by equation (2),

$$CD(oq_i, os_j) = len(chain(oq_i, os_j, SLN_{wiki})) \tag{2}$$

where SLN_{wiki} denotes the semantic link network on Wikipedia, $len()$ denotes the length of a chain of semantic links, $chain(oq_i, os_j, SLN_{wiki})$ denotes the chain of the semantic links connecting oq_i and os_j in SLN_{wiki}. *Category distance* takes an average value because it is hard to distinguish the important semantic objects from the unimportant semantic objects in a question by the semantic link network. Then, the sentences in an answer range can be ranked by *category distance* and the first N sentences are selected as answers.

Actually, answer extracting method incorporates the semantic link network to a simple method that does not have any semantic support. If $D(oq_i, os_j)$ degenerates to 0 - 1 function, which is $D(oq_i, os_j) = 1$ when $oq_i = os_j$ and $D(oq_i, os_j) = 0$ when $oq_i \mathrel{!=} os_j$, the extracting method degenerates to a method that only considers the number of overlapped words between question and answer. So the method removes other influential factors and can reflect the influence of the semantic link network.

We conduct a series of experiments to test the performance of the QA system while changing the scales of given answer range and semantic link network.

6. Dimensions of ability

To further analyze the ability of the semantic link network, we test the ability of the QA system from the following two dimensions [42][51][59].

One dimension is the types of question. We set two groups of question types. One is the group that includes 7 types of question: (1) *Object description*, (2) *Target object*, (3) *Time*, (4) *Location*, (5) *Person*, (6) *Reason*, and (7) *Process*. The group is set based on different types of interrogatives. The details are elaborated in Appendix B. The other group gives a more elaborate division *SMU*. It classifies questions into 16 types: (1) *date*, (2) *definition*, (3) *distance*, (4) *location*, (5) *manner*, (6) *money*, (7) *name*, (8) *NNP*, (9) *number*, (10) *organization*, (11) *person*, (12) *price*, (13) *reason*, (14) *time*, (15) *title*, and (16) *undefined*. The group is set based on the types of information asked by question. We chose *SMU* because *subcategory*, *is-part-of*, and *instance* links can help answer most types of question in *SMU*. Some classifications give more question types than *SMU*. But they contain some question types that the semantic links can hardly help answer, such as *FREQUENCY* or *WEIGHT* in *Cymfony*, or *ABBREV* in *UIUC*.

The other dimension is the patterns of answering. One pattern indicates a way to answer a question and different answers may correspond to different patterns. Specifically, for a question, a pattern indicates a set of answers following same features. For example, for a question "*what is data mining*?", the answer "*data mining is a research field of computer science.*" answers the question by syntactically replacing the interrogative "*what*" with a certain semantic object "*a research field of computer science*", and the answer "*Data mining focuses on discovering patterns in large data sets*" gives the connotation of "*data mining*". The two answers follow different patterns. There may be other answers that meet the two patterns.

Five typical patterns are summarized as follows.

1. *Replacement match*. An answer matches a question if the answer can be obtained by replacing the interrogative in question with one or several semantic objects or a statement.
2. *Common-statement match*. An answer matches a question if the answer and the question contains a common statement and the question asks for complementing the statement.
3. *Semantic-link match*. An answer matches a question if the semantic object in the question and the semantic objects in the answer have some semantic links that meet the way to answer the question.
4. *Sentence-meaning match*. An answer matches a question if the overall meaning of the answer meets the request of the question.
5. *Multi-sentence-meaning match*. An answer matches a question if the answer contains multiple sentences or clauses, which have *sequential* relations and their overall semantics matches the request of the question.

The details of the patterns are introduced in Appendix C.

7. Experiments

7.1. Experiments on Influence

7.1.1. Basic setting and dataset

The experiment tests the performance of the experimental QA system on two changeable factors, answer range and the semantic link network, to reveal the influence between them.

The following are several basic settings:

1. *Answer*. Except the questions and answers that can be directly determined by testers, a sentence can be determined as an answer to a question by the following criteria: (1) A sentence answers an *object description* question if the sentence describes a semantic link about the asked semantic object including subcategories or super-categories, part-of, instances, and connotation. (2) A group of sentences can answer a question if at least one sentence in the group that can answer the question. (3) A group of sentences is considered as an answer to a question on *process* if the sentences cover the steps asked by the question.
2. *Answer range.* It increases with collecting pages from Wikipedia for different experiments. As domain-relevant pages are more likely to contain answers, we set up two comparative groups: (1) Domain-relevant group contains domain-relevant pages or the pages that either are subordinate to a common category or

separately belong to two categories connected by subcategory relation. The pages are on different but relevant topics: *artificial intelligence*, *natural language processing*, *artificial neural network*, *knowledge base*, *machine learning*, and *data mining*. (2) Domain-irrelevant group contains domain irrelevant pages on *data loss*, *natural science*, *veterinary medicine*, *computer science*, *synthetic biology*, and *music psychology*. Experiments are conducted separately on the two groups, each increases from one page to six pages, denoted as P_1, P_2, P_3, P_4, P_5, and P_6. Each page has a page title denoted as K_i (i=1, …, n). Synonyms, alternative names, and abbreviations are added into K_i and linked to the common page of titles.

3. *Semantic link network*. Based on different sets of terms, semantic link networks of different scales are constructed, denoted as $SLN(\{K_1\})$, …, $SLN(\{K_1, K_2, …, K_6\}$. A larger scale of semantic link network contains more semantic links and semantic objects.

4. *Question*. Five testers (graduate students) are invited to ask questions on the pages in the two groups, forming two question sets *QuestionSet*$_1$ and *QuestionSet*$_2$. To conduct targeted experiments on object description, we select the type of questions from *QuestionSet*$_1$ and *QuestionSet*$_2$ respectively to form two other question sets: *ODQuestionSet*1 and *ODQuestionSet*$_2$.

5. *Dataset*. The experiment sets a series of QA tasks. Each task consists of two steps: (1) fix an answer range, a semantic link network, and a question set; and, (2) use the experimental QA system to extract answers from the answer range for the questions in the question set based on the semantic link network. One group of texts has 6 answer ranges, 6 corresponding semantic link networks, and 4 question sets. Their combination generates 144 QA tasks for one group (totally 288 tasks for two groups). The data and the results can be downloaded at https://ln2.sync.com/dl/37827f3d0/k2q3q7jf-fvzmsyb4-b75g9vxc-fy6ck2r3/view/default/8782970980006.

7.1.2. Results

Fig. 3 shows the results of the experiments on *QuestionSet*$_1$ and *ODQuestionSet*$_1$. The two horizontal axes represent the answer range and the semantic link network. The vertical axis represents the correct ratio calculated by the number of the questions that are correctly answered / the total number of the questions. Fig. 3 indicates the following points.

1. The precisions in (a) and (b) increase with widening the answer range and expanding the semantic link network, which verifies the overall influence conjecture.
2. In (a) and (b), the precisions increase with widening the answer range while fixing a point at the axes of semantic link network, and the precisions increase with expanding the semantic link network while fixing a point at the axes of answer range. The results verify the side influence conjecture.

To show the mutual-influence between answer range and the semantic links, we highlight the relatively medium precisions in Fig. 3 (a) and (b). The two figures are transformed to Fig. 4 (a) and (b) by keeping the precisions around 50%-60%. In Fig. 4 (a) and (b), the black cubes represent the positions of the kept precisions in Fig. 3 (a) and (b). Fig. 4 (a) and (b) show that the kept precisions are distributed along with the diagonal lines from $(K_1, P_1 \cup P_2 \cup P_3 \cup P_4 \cup P_5 \cup P_6)$ to $(K_1 \cup K_2 \cup K_3 \cup K_4 \cup K_5 \cup K_6, P_1)$. So, to maintain the precision at a certain range, narrowing the answer range requires adding more semantic links and removing the semantic links implies to widen the answer range. This verifies the mutual influence proposition.
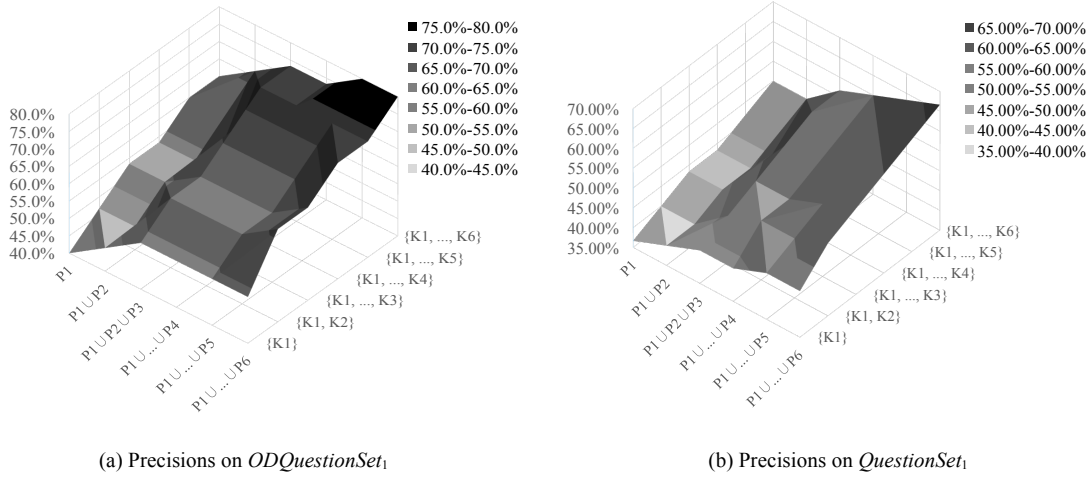
(a) Precisions on *ODQuestionSet₁*

(b) Precisions on *QuestionSet₁*

Fig. 3. The results of performing the QA system on the domain-relevant group. (a) The results of answering the questions in *ODQuestion-Set₁*. (b) The results of answering the questions in *QuestionSet₁*.
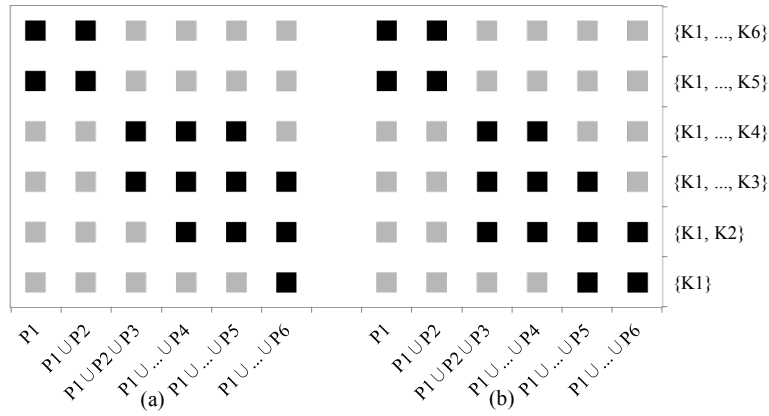

Fig. 4. The transformations of Fig. 3 (a) and (b). (a) The precisions around 50%-60% on *ODQuestionSet₁*. (b) The precisions around 50%-60% on *QuestionSet₁*.

Fig. 5 (a) and (b) show the results of the experiments on *QuestionSet₂* and *ODQuestionSet₂*. Because the pages are domain irrelevant, each time of widening the answer range provides more answers corresponding to only a part of the questions by chance; and each time of expanding the semantic link network covers more semantic objects for only a part of the questions by chance as well. Fig. 5 (a) and (b) indicate the following points.

1. Fixing a point at the axes of answer range in (a) and (b), the precisions almost keep unchanged with the expansion of semantic link network because there are insufficient answers. So, the side influence proposition and mutual influence proposition are not valid when the potential answer sentences are insufficient.

2. Fixing a point at the axes of semantic link network in (a) and (b), the precisions increase with expanding answer range. Especially, the precisions still increase when the semantic link networks do not cover the semantic objects in answer range. For example, when the semantic link network is $SLN(\{K_1\})$, only the semantic objects and links in P_1 are involved, but the precision still increases with widening the answer ranges. It indicates that providing more texts can improve the performance of the QA system even though

the added texts are not covered by the semantic link network. So the answer range plays a more important role than the semantic link network for the QA system.
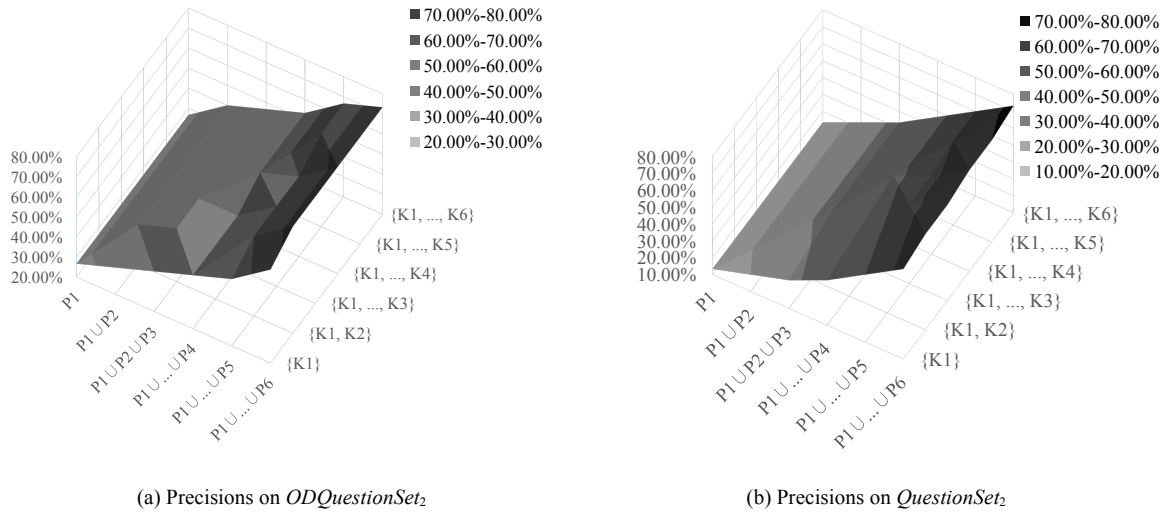

(a) Precisions on *ODQuestionSet₂*

(b) Precisions on *QuestionSet₂*

Fig. 5. The results of performing the QA system on irrelevant-group. (a) The results of answering the questions in DQS_2. (b) The results of answering the questions in QS_2.

7.2. Experiments on question type

7.2.1. Basic setting and dataset

The experiments test the QA system on answering different types of questions and supporting different patterns. Because the ability of classification is prominent when a question concerns specialized knowledge, the experiment uses academic papers and Wikipedia pages as answer range. To reflect the improvement by using the semantic link network, we use a comparable answer extraction approach that does not use any semantic information. The approach calculates the cosine distance between a question and the sentences in answer range.

The dataset contains 100 academic papers from *AAAI* conference and 1100 pages from Wikipedia. Five testers are invited to throw questions on each paper or page. Each question belongs to the 7 types and the *SMU*. The dataset contains 4975 questions, 6991 answers extracted by the QA system and 3801 answers extracted by cosine approach. The data and the results are available for download at https://ln2.sync.com/dl/37827f3d0/k2q3q7jf-fvzmsyb4-b75g9vxc-fy6ck2r3/view/default/8782970980006.

7.2.2. Results

First, the experiment tests the QA system's performance on answering the 7 types of question (totally 4975 questions). For each time of question answering, 5 sentences are extracted as the answers.

Table 1. The experimental results on the 7 types of questions.

Question type	No. of CRQ by SLN	CR	No. of CRQ by co-sine	CR	IM	No. of CS by SLN	CR	No. of CS by cosine	CR	IM
Object description	1502/2261	0.66	655/2261	0.29	**0.37**	3795/11305	0.34	1441/11305	0.13	**0.21**
Time	317/405	0.78	147/405	0.36	**0.42**	423/2025	0.21	243/2025	0.12	**0.09**
Person	282/373	0.76	161/373	0.43	**0.33**	482/1865	0.26	223/1865	0.19	**0.07**
Location	111/133	0.84	32/133	0.24	**0.60**	178/665	0.27	99/665	0.15	**0.12**
Target object	618/1080	0.57	498/1080	0.46	**0.11**	2054/5400	0.38	1727/5400	0.32	**0.06**
Reason	25/254	0.1	19/254	0.07	**0.03**	37/1270	0.03	50/1270	0.04	**-0.01**
Process	15/469	0.03	11/469	0.02	**0.01**	22/2345	0.01	18/2345	0.01	**0**
Average CR	Total: 2870/4975	0.58	Total: 1523/4975	0.31	**0.27**	Total: 6991/24875	0.28	Total: 3801/24875	0.15	**0.13**

- SLN denotes the experimental QA system that uses the semantic link network.
- CRQ denotes correctly answered questions.
- CR denotes correct ratio.
- CS denotes the correct sentences (answers) to a question.
- IM denotes the improvement that equals to the correct ratio by SLN minus the correct ratio by cosine approach.
- The value a/b in columns "No. of CRQ by SLN" and "No. of CRQ by cosine" denotes (the number of CRQ) / (total number of question).
- The value a/b in columns "No. of CS by SLN" and "No. of CS by cosine" denotes (the number of CS) / (total number of extracted sentences).

Table 2. The experimental results of the 4 forms of answering object description questions.

Object description	No. of CS	Sub or super-category	Part	Instance	Connotation
SLN	3795	1727	1176	624	268
CR	/	0.46	0.31	0.16	0.07
cosine	1441	331	461	275	374
CR	/	0.23	0.32	0.19	0.26

Table 1 shows the following points:

1. The experimental QA system can approximately answer 2/3 of the *object description* questions while cosine approach can nearly answer 1/3. This indicates that the semantic link network can significantly help answer the type of question. More specifically, there are 4 ways to answer the *object description* questions: subcategory or super-category, part-of, instance, and connotation. We classify the 3795 correct sentences by the QA system and the 1441 correct sentences by cosine approach into the 4 forms as shown in Table 2. This shows that (1) the system generates the correct sentences (answers) of the *object description* questions that mostly belong to the first 3 forms and rarely belong to the connotation form, and (2) by using the cosine approach, the correct sentences are relatively evenly distributed in the 4 forms. This indicates that, for *object description* question, the semantic link network can help extract answers in the forms of sub-category or super-category, part and instance.

2. The performance of answering the *person*, *time* and *location* questions is relatively good and the improvements on the 3 types are relatively great. There are two reasons: (1) filtering function removes most sentences that mismatch the type of the question; and, (2) the semantic objects in some answers happen to have the semantic links connected to the semantic objects in question. For example, for the question "*When is decision tree used in NLP?*", the sentence "*Starting in the late 1980s, however, there was a revolution in NLP with the introduction of machine learning algorithms for language processing.*" is extracted because there are *is-part-of* semantic links: *decision trees → classification algorithms → machine*

learning. However, the second situation rarely happens. On the other hand, the correct ratios of the correct sentences are relatively low. This is because the 3 types of questions correspond to a few answers, which are unlike *object description* question that may correspond to multiple acceptable answers.

3.   The performance of answering *target object* questions is medium and the improvement is little. The reason is that there are mainly 4 kinds of *target object* questions according to the way of an answer giving target objects, and the semantic link network can only help a part of them. The 4 kinds include: (a) Giving an attribute, for example, question "*What sorting algorithm is iterative*?" where "*iterative*" is an attribute of the target object, (b) giving a function, for example, question "*What NLP tool can do named entity recognition from texts*?" where "*named entity recognition*" is a function of the target object, (c) giving the semantic links, for example, for the question "*what are the sub-fields of artificial intelligence research*?" where "*Artificial intelligence*" has a semantic link "*subcategory*" to the target object of the question, (d) giving a statement, for example, for the question "*What software does a computer use to store data*?", "*software*", "*computer*", "*data*" and the target object "*database management system*" form a statement that is "*Computers use database management system to store data.*". The given semantic link networks can hardly help the answers that give attribute, function and statement.

4.   The performance of answering *reason* and *process* questions is relatively weak because the semantic links do not contain *cause-effect* link and *sequential* link. However, there are a few *reason* and *process* questions that are correctly answered because the questions happen to the concern on the semantic links on category. For example, for the question "*Why is bioinformatics an interdisciplinary subject*?", the semantic objects in the answer "*Bioinformatics uses the algorithms of computer science to deal with the data from biology.*" have the semantic links "*subcategory*" to "*bioinformatics*". Therefore, the performance of the QA system on the two types is a little better than the performance of using the cosine approach.

Second, to make a solid result, we further conduct an experiment on question type *SMU*. Because people seldom ask *distance*, *money*, *price* and *title* questions on the texts describing specialized knowledge and the semantic link on category can hardly help answer the questions, the 4 types are neglected. So there are 12 types left. Five testers (graduate students) are required to ask questions (total 4244 questions) on each text in the dataset. Table 3 shows the results that indicate the following points.

1.   In general, the semantic link network can help extract more answers.
2.   The performance of answering *data*, *time*, *person*, *location*, *name*, and *NNP* questions by the QA system are relatively good and the improvement on them is relatively great. This is because they benefit from filtering function. The results are consistent with the results of answering *time*, *person* and *location* questions in the 7 types, which benefit from filtering function as well. However, filtering function sometimes cannot help the types of questions. For example, to answer a *time* questions "*when did the project of the fifth generation computer system fail*?", a QA system needs to know more knowledge to distinguish the starting time point and the terminating time point.
3.   The performance of *manner* and *reason* questions by the QA system is relatively weak and the improvement on them is weak, especially for *reason* questions. The results are consistent with the results on *reason* and *process* questions in 7 types.
4.   The performance of answering the *definition* questions by the QA system is relatively good and the improvement is relatively great, which is consistent with the results on *object description* question in the 7 types.
5.   Even though filtering function can help answer *organization* and *number* question, the performance of the QA system and the improvement on the types are weak. This is because a lot of sentences contain numbers which weakens filtering function, some *number* questions require inferential knowledge to do a simple math, and the semantic links on category only appear between the asked organization and the objects in question.

Table 3. The experimental result on SMU.

Question type	Correct/ Total by SLN	CR	Correct/ Total by cosine	CR	IM	Correct/ Total by SLN
date	177/221	0.80	75/221	0.34	0.46	177/221
time	153/200	0.77	58/200	0.29	0.48	153/200
person	440/593	0.74	196/593	0.33	0.41	440/593
location	71/100	0.71	21/100	0.21	0.5	71/100
definition	573/816	0.70	277/816	0.34	0.36	573/816
name	244/379	0.64	110/379	0.29	0.35	244/379
unidentified	64/100	0.64	47/100	0.47	0.17	64/100
NNP	197/314	0.63	97/314	0.31	0.32	197/314
manner	153/325	0.47	66/325	0.2	0.27	153/325
organization	38/87	0.44	16/87	0.18	0.26	38/87
reason	202/748	0.27	165/748	0.22	0.05	202/748
number	84/361	0.23	108/361	0.3	-0.07	84/361
Total	2396/4244	0.56	1236/4244	0.29	0.27	2396/4244

7.3. Experiment on pattern

7.3.1. Basic setting and dataset

The experiment tests the abilities of the QA system and the cosine approach on the 5 patterns. We use the dataset in the experiment on question type. 4975 questions, 6991 correct sentences (answers) generated by the QA system, and 3801 correct sentences from cosine approach are classified into the 7 patterns. One question and one answer may match multiple patterns from the testers' points of view because they do not know how the answers are extracted. So we stipulate that the testers can only choose one pattern that matches most closely for a pair of question and answer.

7.3.2. Results

Table 4 shows the distributions of the answers on the 7 question types.

Table 4. The distributions of the answers in the 5 patterns.

Question type	By SLN					By cosine				
	No. of answers	RM	CSM	SLM	SMM	MSMM	No. of answers	RM	CSM	SLM	SMM	MSMM
Object description	3795	2239	0	1283	273	0	1441	1031	0	327	83	0
Time	423	113	279	31	0	0	243	134	35	53	21	0
Person	482	173	236	73	0	0	223	95	16	63	49	0
Location	178	56	100	22	0	0	99	46	8	27	18	0
Target object	2054	727	275	914	138	0	1727	523	480	313	411	0
Reason	37	19	0	18	0	0	50	11	0	8	31	0
Process	22	15	0	0	2	5	18	7	0	2	1	8
Total	6991	3342	890	2341	413	5	3801	1847	539	793	614	8

Fig. 6 and Fig. 7 are derived from Table 4 through normalization. The two figures show the distributions of the answers in the 5 patterns intuitively. The parts "*SUM*" in the two figures show the total distributions. RM denotes replacement match pattern; CSM denotes common statement match pattern; SLM denotes semantic-link match pattern; SMM denotes sentence meaning match pattern; and MSMM denotes multi-sentence meaning match pattern.
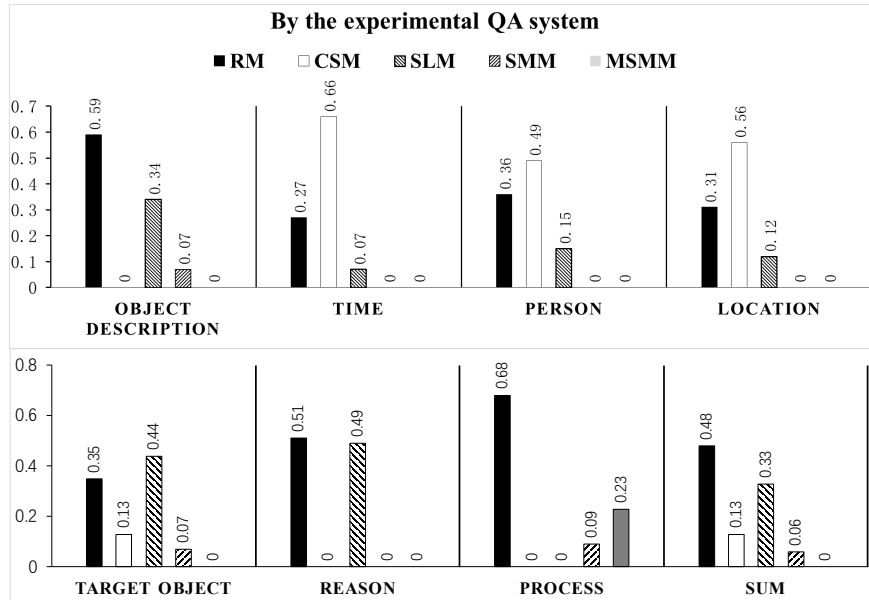

Fig. 6. The distribution of the correct sentences from the experimental QA system on the 5 patterns.
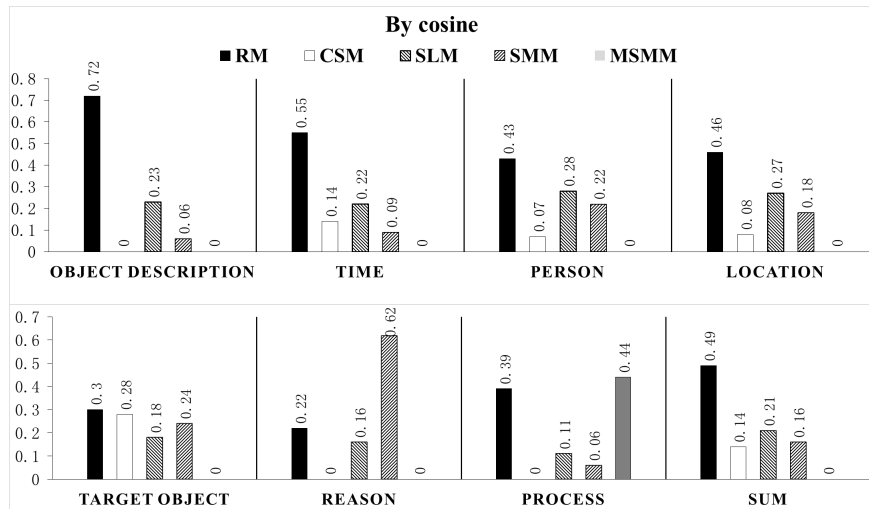

Fig. 7. The distribution of the correct sentences from cosine approach on the 5 patterns.

Table 4, Fig. 6 and Fig. 7 indicate the following points.

(1) By using the QA system, Table 4 shows that a lot of answers and questions match replacement match pattern, and the ratios of the pattern in Fig. 6 are relatively high, especially for *object description* question, and the ratios of the pattern using cosine is also high. The reason is that the pattern is the most common way for people to identify answer. However, the number of the answers matching the pattern by the QA system is much greater than by cosine approach. The reason is that the semantic links can help identify semantically related semantic objects of the semantic objects in the questions. For example, an answer to the question "what is natural language processing?" sometimes can also be considered as an answer to the question "What is artificial intelligence?". Therefore, the semantic link network can help match replacement match pattern.

(2) Table 4 shows that, with common-statement match pattern, the answers from the QA system are much more than the answers from the cosine approach. This indicates that the semantic link network can help the pattern. Furthermore, most of the exceeding answers come from *time*, *person* and *location* questions. This indicates that the improvement is mostly based on filtering function.

(3) Table 4 shows that, with the semantic-link match pattern, the answers from the QA system are much more than the answers from the cosine approach. It indicates that the semantic link network can help the semantic-link match pattern. On the other hand, the ratio (0.33) of the pattern in part "*SUM*" of Fig. 6 is greater than the ratio (0.21) of the pattern in part "*SUM*" of Fig. 7. This indicates that the semantic link network helps the pattern more greatly than other patterns. This is because the QA system extracts answers by tracing semantic link. However, the QA system only uses the three types of semantic link. So the QA system can hardly extract answer concerning other types of semantic links.

(4) A relatively small number of answers and questions matches sentence-meaning match pattern, both by the QA system and cosine approach. The cosine approach is better than the QA system in numbers and ratios. This is because the pattern usually concerns other types of semantic links. So the three types of semantic link can help the pattern to a certain extent.

(5) A few answers and questions match the multi-sentence meaning match pattern. This is because the pattern is mostly used for *process* questions that need the answers containing multiple sequential actions and *reason* questions the need multiple steps of inferring. These questions and answers usually concerns other types of semantic links. So the three types of semantic link can support the pattern to a certain extent.

In general, the three types of semantic link can help replacement match pattern, common-statement match pattern and semantic-link match pattern, but they can hardly help sentence meaning match pattern and multi-sentence meaning match pattern (more types of semantic links are needed). Compared with other patterns, the semantic link network helps semantic-link match pattern more strongly.
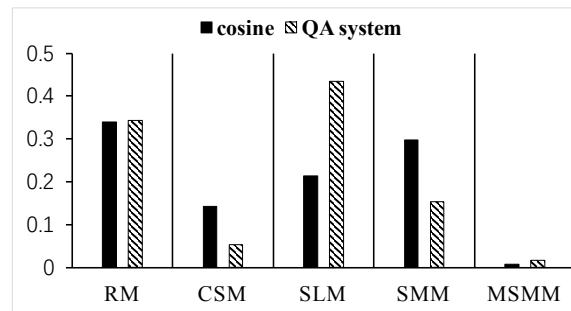

Fig. 8. The distribution of the answers on SMU on the 5 patterns.

To further verify the results, we classify the answers on *SMU* (2396 answers from the QA system and 1236 answers from cosine approach) into the 5 patterns. In general, more answers are extracted by the QA system

compared with cosine approach, and the semantic link network can support the patterns to different extent as shown in Fig. 8. In semantic-link match pattern, the ratio of the QA system is much greater than the ratio of cosine approach, and the difference between the two ratios is greater than the difference with other patterns. This indicates that the semantic link network supports semantic-link match pattern more strongly than other patterns. Therefore, the results in Fig. 8 are consistent with the results as shown in Fig. 6 and Fig. 7.

8. Conclusions

The performance of a QA system can be improved by providing more relevant texts or a larger-scale of semantic link network on texts. Fixing an answer range and a question, a QA system is more likely to find more answers by providing richer semantic links on the answer range; and fixing a question and the scale of semantic link network, more answers can be extracted by adding more relevant texts to the answer range. Keeping a certain range of performance leads to the following results: (a) a narrower answer range implies that the QA system needs a larger scale semantic link network; and, (b) a smaller scale of semantic link network implies that the QA system needs a wider answer range. The side influence and mutual influence is not obvious when an answer range contains insufficient answer sentences, even though a larger scale of semantic link network is provided. So the answer range influences the performance of the QA system more significantly than the semantic links as this type of QA system relies on extracting sentences.

In addition, the experiments show that the three types of semantic link have the following effect: (1) they can help answer *object description* questions when an answer describes the subcategory, super category, part-of, and instance of the asked semantic object, (2) they can hardly help extract answers that contain the connotation of asked semantic object, (3) they can help answer *time*, *person*, and *location* questions because of the filtering function, (4) they can hardly help answer *reason* and *process* questions because they cannot represent *cause-effect* and *process* relations, except that the questions happen to concern the three types of semantic link. It is desirable to be helpful if the *cause-effect* link, *similar* link, and *sequential* link are incorporated into the QA system [52, 53].

The semantic link network helps the replacement match pattern by identifying semantically-matched semantic objects. It helps common statement match pattern with the filtering function. It supports semantic-link match pattern when the question concerns the three types of semantic link. Although it can hardly help sentence meaning match pattern and multi-sentence meaning match pattern, it is desirable to be helpful when more types of semantic links such as *cause-effect*, *similar*, and *sequential* links are employed in the system. This also verifies the importance of the diversity of semantic links [44][45][46].

In short, incorporating semantic link network into a QA system can enhance its performance in answering the questions covered by the answer range and the richness of semantic links.

Acknowledgement

References

1. Adafre, S. F., Jijkoun, V., and Rijke, M. D. Link-Based vs. Content-based retrieval for question answering using Wikipedia. *Evaluation of Multilingual and Multi-modal Information Retrieval*. Springer Berlin Heidelberg, 537-540, 2006.
2. Ahn, D., Jijkoun, V., Mishne, G., Müller, K., Rijke, M., and Schlobach, S. Using Wikipedia at the TREC QA track. *Proceedings of the Text Retrieval Conference* (*TREC*), 2004.
3. Zhou, G. D., & Su, J. Named entity recognition using an HMM-based chunk tagger. *Meeting on Association for Computational Linguistics*, 473-480, Association for Computational Linguistics, 2002.
4. Bollacker, K., Evans, C., Paritosh, P., Sturge, T., and Taylor, J. Free-base: a collaboratively created graph database for structuring human knowledge. *Proceedings of the 2008 ACM SIGMOD International Conference on Management of Data*, Vancouver, Canada, 1247-1250, 2008.

5. Burke, R. D., Hammond, K. J., Kulyukin, V., Lytinen, S. L., Tomuro, N., and Schoenberg, S. Question answering from frequently asked question files: Experiences with the FAQ finder system. *AI magazine*, 18(2), 57, 1997.

6. Woods, W. A., and Kaplan, R. Lunar rocks in natural English: Explorations in natural language question answering. *Linguistic structures processing*, 5, 521-569, 1977.

7. Wang, B., Liu, B., Wang, X., Sun, C., & Zhang, D. Deep learning approaches to semantic relevance modeling for Chinese question-answer pairs. *ACM Transactions on Asian Language Information Processing*, 10(4), 1-16, 2011.

8. Clarke, C., Cormack, G., Kisman, D., and Lynam, T. Question answering by passage selection (multitext experiments for TREC-9). *Proceedings of the 9th Text Retrieval Conference* (*TREC*), 2000.

9. Cui, H., Kan, M. Y., and Chua, T. S. Unsupervised learning of soft patterns for generating definitions from online news. *Proceedings of the 13th International Conference on World Wide Web* (*WWW*), New York, NY, USA, ACM, 90-99, 2004.

10. Cui, H., Sun, R., Li, K., Kan, M., and Chua, T. Question answering passage retrieval using dependency relations. *Proceedings of the 28th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval* (*SIGIR*), New York, NY, USA: ACM, 400-407, 2005.

11. Dan, M., Harabagiu, A., Girju, R., Morarescu, P., Lacatusu, F., Novischi, A., Badulescu, A., and Bolohan, O. LCC tools for question answering. *Proceedings of the 11th Text Retrieval Conference* (*TREC*), 2002.

12. Dan, M., Paşca, M., Harabagiu, S., and Surdeanu, M. Performance issues and error analysis in an open-domain question answering system. *ACM Transactions on Information Systems*, 21(2), 133-154, 2003.

13. Duan, H., Cao, Y., Lin, C., and Yu, Y. Searching questions by identifying question topic and question focus. *Proceedings of the Meeting of the Association for Computational Linguistics* (*ACL*), 156-164, 2008.

14. Echihabi, A., and Marcu, D. A noisy-channel approach to question answering. *Proceedings of the 41st Annual Meeting on Association for Computational Linguistics* (*ACL*), Stroudsburg, PA, USA, Association for Computational Linguistics, 16-23, 2003.

15. Ganesh, S., and Varma, V. Exploiting structure and content of Wikipedia for Query Expansion in the context of Question Answering. *Proceedings of the international conference on recent advances in natural language processing* (*RANLP*), 103-106, 2009.

16. Gobeill, J., Pasche, E., Teodoro, D., Veuthey, A. L., and Ruch, P. Answering gene ontology terms to proteomics questions by supervised macro reading in medline. *Embnet Journal*, 18(B), 29-31, 2012.

17. Green, Jr B. F., Wolf, A. K., Chomsky, C and Laughery, K. Baseball: an automatic question answerer, *Proceedings of western joint IRE-AIEE-ACM computer conference*, 219-224, 1961.

18. Hovy, E., Gerber, L., Hermjakob, U., Junk, M., and Lin, C. Y. Question Answering in Webclopedia. *Proceedings of the Text Retrieval Conference* (*TREC-9*), 655-664, 2001.

19. Feng, M., Xiang, B., Glass, M. R., Wang, L., and Zhou, B. Applying deep learning to answer selection: A study and an open task. *Automatic Speech Recognition and Understanding*, 813-820, 2016

20. Ko, J., Nyberg, E., and Luo, S. A probabilistic graphical model for joint answer ranking in question answering. *Proceedings of the, International ACM SIGIR Conference on Research and Development in Information Retrieval* (*SIGIR*), Amsterdam, the Netherlands, 343-350, 2007.

21. Lee, G. G., Seo, J., Lee, S., Jung, H., Cho, B., Lee, C., Kwak, B., Cha, J., Kim, D., An,. J., Kim, H., and Kim, K. SiteQ: engineering high performance QA system using lexico-semantic pattern matching and shallow NLP. *Proceedings of the 10th Text Retrieval Conference* (*TREC*), 2001.

22. Lehmann, J., Isele, R., Jakob, M., Jentzsch, A., Kontokostas, D., Mendes, P. N., Hellmann, S., Morsey, M., van Kleef, P., Auer, S., Bizer, C. Dbpedia - A large-scale, multilingual knowledge base extracted from Wikipedia. *Semantic Web*, 6(2), 167-195, 2015.

23. Leidner, J., Bos, J., Dalmas, T., Curran, J., Clark, S., Bannard, C., Webber, B., and Steedman, M. Qed: The edinburgh trec-2003 question answeringsystem. *Proceedings of the Text REtrieval Conference* (*TREC*), 631-635, 2003.

24. Wang, C, Xiong, M., Zhou, Q., Yu, Y. PANTO: A Portable Natural Language Interface to Ontologies. *The Semantic Web: Research and Applications*, 473-487, 2007.

25. Dan, M., Harabagiu, S., Pasca, M., Mihalcea, R., Girju, R., Goodrum, R., and Rus, V., The Structure and Performance of an Open-Domain Question Answering System. *Proceedings of the 38th Annual Meeting of the Association for Computational Linguistics* (*ACL*), October 2000.

26. Lita, L. V., Hunt, W. A., and Nyberg, E. Resource analysis for question answering. *Proceedings of the 42th Annual Meeting on Association for Computational Linguistics* (*ACL*), 162-165, ACL, 2004.

27. Li, X., and Roth, D. Learning question classifiers. *Proceedings of the 19th International Conference on Computational Linguistics* (*Coling*), Stroudsburg, PA, USA, Association for Computational Linguistics, 1-7, 2002.

28. McGuinness, D. L.，and Van, H. F. OWL web ontology language overview. *W3C recommendation*, 10, 135 -140, 2004 .

29. Guo, Q., and Zhang, M. Question Answering Based on Pervasive Agent Ontology and Semantic Web. *Knowledge-Based Systems*，

22, 443-448, 2009.

30. Tablan, V, Damljanovic, D., and Bontcheva, K. A Natural Language Query Interface to Structured Information. *Proceedings of the 5th European Semantic Web Conference on the Semantic Web: Research and Applications* (*ESWC*), 2008.

31. Moldovan, D. I., and Rus, V. Logic form transformation of wordnet and its applicability to question answering. *Proceedings of the 39th Annual Meeting on Association for Computational Linguistics* (*ACL*), 402-409, 2001.

32. Nastase, V., and Strube, M. Decoding Wikipedia categories for knowledge acquisition. *Proceedings of the 23rd National Conference on Artificial Intelligence* (*AAAI*), 1219-1224, 2008.

33. Zhuge, H. The Knowledge Grid, World Scientific Publishing Co., 2004.

34. Manning, C. D., Surdeanu, M., Bauer, J., Finkel, J., Bethard, J. S., and McClosky, D. The Stanford CoreNLP Natural Language Processing Toolkit. *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics* (*ACL*), 55-60, 2014.

35. Smith, B., Ashburner, M., Rosse, C., Bard, J., Bug, W., Ceusters, W., Goldberg, L. J., Eilbeck, K., Ireland, A., Mungall, C. J., Leontis, N., Rocca-Serra, P., Ruttenberg, A., Sansone, S-A., Scheuermann, R. H., Shah, N., Whetzel, P. L., and Lewis, S. The OBO Foundry: coordinated evolution of ontologies to support biomedical data integration. *Nature biotechnology*, 25(11):1251-1255, 2007.

36. Zhang, D., Lee, W. S. Question classification using support vector machines. *Proceedings of the 26th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval* (*SIGIR*), New York, NY, USA, ACM, 26-32, 2003.

37. Bollacker, K., Evans, C., Paritosh, P., Sturge, T., and Taylor, J. Freebase: a collaboratively created graph database for structuring human knowledge. *Proceedings of the 2008 ACM SIGMOD international conference on Management of data*, 1247-1250, 2008.

38. Tellex, S., Katz, B., Lin J, Fernandes, A., and Marton, G. Quantitative evaluation of passage retrieval algorithms for question answering. *Proceedings of the 26th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval* (*SIGIR*), New York, NY, USA, ACM, 41-47, 2003.

39. Chen, D., Fisch, A., Weston, J., and Bordes, A. 2017. Reading Wikipedia to answer open domain questions. arXiv preprint arXiv:1704.00051, 2017.

40. Wilensky, R., Chin, D. N., Luria, M., Martin, J., Mayfield, J., and Wu, D. The Berkeley UNIX consultant project. *Intelligent Help Systems for UNIX*, Springer Netherlands, 49-94, 2000.

41. Xue, X., Jeon, J., and Croft, W. B. Retrieval models for question and answer archives. *Proceedings of the 31st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval* (*SIGIR*), New York, NY, USA, ACM, 475-482, 2008.

42. Zhuge, H. The Web Resource Space Model, *Springer*, 2018.

43. Banerjee, S., and Pedersen, T. Extended gloss overlaps as a measure of semantic relatedness. *Proceedings of the 18th International Joint Conference on Artificial Intelligence* (IJCAI), 805-810, 2003.

44. Zhuge, H. Communities and Emerging Semantics in Semantic Link Network: Discovery and Learning. *IEEE Transactions on Knowledge and Data Engineering*, 21(6), 785-799, 2009.

45. Zhuge, H. Interactive Semantics, *Artificial Intelligence*, 174, 190-204, 2010.

46. Zhuge, H. Semantic Linking through Spaces for Cyber Physical Socio Intelligence: A Methodology, *Artificial Intelligence*, 175, 988-1019, 2011.

47. Zhuge, H. Multi-Dimensional Summarization in Cyber-Physical Society, Morgan Kauffman, 2016.

48. Zhuge, H. The Knowledge Grid: Toward Cyber-Physical Society. 2nd edition. Singapore: World Scientific Publishing Co.; 2012.

49. Zhuge, H. Inheritance rules for flexible model retrieval. Decision Support Systems. 22(4), 379-90, 1998.

50. Zhuge, H. Active e-document framework ADF: Model and tool. Information & Management. 41(1), 87-97, 2003.

51. Zhuge, H. and Xing, Y. Probabilistic Resource Space Model for Managing Resources in Cyber-Physical Society. *IEEE Transactions on Services Computing*, 5(3), 404-21, 2012.

52. Cao, M., Sun, X., and Zhuge, H. The Contribution of Cause-Effect Link to Representing the Core of Scientific Paper - The Role of Semantic Link Network. PLOS One, vol.13, no.6, 2018.

53. Sun, X. and Zhuge, H. Summarization of Scientific Paper through Reinforcement Ranking on Semantic Link Network, *IEEE Access*, vol. 6, 2018, p.40611-40625.

54. Zhou, L., Gao, J., Li, D., Shum, H-Y., The Design and Implementation of Xiaoice, an Empathetic Social Chabot, https://arxiv.org/PDF/1812.08989v1.pdf

55. Zhuge, H. Cyber-Physical-Social Intelligence on Human-Machine-Nature Symbiosis, Springer, 2020.

56. Hill, J., Ford, W. R., Farreras, I. G. Real conversations with artificial intelligence: A comparison between human–human online conversations and human–chatbot conversations. Computers in Human Behavior, 49, 245-250, 2015.

57. Shang, L., Lu Z. and Li, H. Neural responding machine for short-text conversation. *Proceedings of Annual Meeting of the Association for Computational Linguistics* (ACL), 1577–1586, 2015

58. Hovy, E., Hermjakob, U., Lin, C.Y., Ravichandran, D. Using Knowledge to Facilitate Factoid Answer Pinpointing, *Proceedings of*

the COLING-02, 2002.

59. Zhuge, H., Xing, Y. and Shi, P. Resource Space Model, OWL and Database: Mapping and Integration. ACM Transactions on Internet Technology 8(4)(2008).

Biography

Bei Xu is an associate professor in Nanjing University of Posts and Telecommunications, China. He received PhD degree from the University of Chinese Academy of Sciences.

Hai Zhuge is a Distinguished Scientist of the ACM, a Fellow of British Computer Society, and a Senior Member of IEEE. He has made systematic contribution to semantics and knowledge modelling through lasting fundamental research on the Semantic Link Network and the Resource Space Model based on a multi-dimensional methodology. He has also made fundamental contribution to the study of Cyber-Physical-Social Intelligence. As a Chair in Computer Science at Aston University, he leads the International Research Network on Cyber-Physical-Social Intelligence consisting of Guangzhou University, Key Lab of Intelligent Information Processing at the Institute of Computing Technology in Chinese Academy of Sciences, Aston University, and University of Chinese Academy of Sciences. He was an associate Editor-in-Chief of Future Generation Computer Systems. He is serving as an Associate Editor of IEEE Intelligent Systems.   Personal webpage: http://www.knowledgegrid.net/~h.zhuge.

Appendix A. Components of semantic ink network

The QA system sets an answer range represented as a set $AR = \{s_i \mid s_i$ is a sentence$\}$. Any free-text based QA system stipulates an answer range explicitly or implicitly for each task of extracting answer. For example, some QA systems collect as many texts as possible so that the answer range is enlarged and therefore the systems can answer more questions, and some QA systems fix answer range within one text or one kind of texts in particular domain.

As some categories or instances in Wikipedia do not indicate semantic objects, they are ignored by the QA system with the following rules: (1) the first letter is not a capital one, (2) the title is a single stop word, (3) for a multiword title, not all words other than prepositions, determiners, conjunctions, or negations are capitalized, and (4) the title has more than four words. As Wikipedia does not contain the terms on some specific things, the QA system uses Named Entity Recognition (Stanford *NLP* toolkit, https://nlp.stanford.edu/software/) to provide a certain types of semantic objects [3]. Specifically, a semantic object is represented by $so = \{s, Type\}$ where s is a string and Type = {time | character | location}.

A semantic link is represented by a triple $sl = <o_1, o_2, link>$ where o_1 and o_2 are two semantic objects and *link* denotes the type of the semantic link. A semantic link network abides by the following rules.

Rule 1: Any semantic link contains category "*main topic classifications*" because a systematic classification always contains a semantic object referring to the range of everything. In this case, "*main topic classifications*" refers to the object. This rule guarantees that the knowledge base is complete from the top.

Rule 2: Except *main topic classifications*, any semantic object has at least one superior semantic object because the meaning of a semantic object implies the meaning of the superior semantic object. If a semantic object has no superior object, a semantic link is established between the semantic object and *main topic classifications*. The rule guarantees that there is no dead end while tracing semantic links until reaching the *main topic classifications*.

Rule 3. Semantic links on classification are based on the relations between Wikipedia's page and category or by rule 2. This rule guarantees that personalized perspectives on classification are unified.

Appendix B. Types of questions

1. *Object description*. This type of questions asks for the description of a semantic object. The basic form is "*what is/are + a term*", e.g., "*What is semantic network?*" and "*What is cyber space?*" [46].

2. *Target object*. This type of questions asks for one or several specific semantic objects or semantic links. The basic form is "*what + a description on one or several semantic objects or semantic links*", e.g., "*What are the links between documents in the World Wide Web*", "*What is the fusion of the physical space, the cyber space, and the socio space?*" and "*What are the basic semantic links in the physical space?*" [46].

3. *Time*. The type of questions asks about time. The interrogatives include "*when*", "*what is the time of*" and "*which/what + word about time*", e.g., "*which month*", etc. Most questions on *Time* follow the two forms: (a) "*interrogative + a term indicating a time*", e.g., "*When is the landmark year for artificial intelligence?*" (from the Wikipedia page on AI), and (b) "*interrogative + an event*", e.g., "*When did Turing conceive the modern computer?*" [46].

4. *Location*. This type of questions asks about location. The interrogatives include "*where*", "*which + a term that means a kind of place*", e.g., "*which city*", and "*what is the place/location/...*", etc. Most *location* questions follow the following two forms: (a) "*interrogative + a term indicating a location*", e.g., "*Where is Carnegie Mellon University?*" (from the Wikipedia page on AI), and (b) "*interrogative +an event*", e.g., "*Where was AI research born at?*" (from the Wikipedia page on AI).

5. *Person*. This type of questions asks about *person*. The interrogatives include "*Who...*" and "*which + a term indicating person*", e.g., "*which expert*", etc. Most *person* questions follow the following two forms: (a) "*interrogative + a term indicating a person*", e.g., "*Who is Vannevar Bush?*", and (b) "*interrogative + a description or an action of a person*", e.g., "*Who conceive the modern compute?*".

6. *Reason*. This type of questions asks about *reason*. The interrogatives include "*why…*" and "*what is/are the reason(s)…*", etc. Most *reason* questions follow the two forms: (a) "*interrogative + a term indicating a phenomenon or an event*", e.g., "*what is the reason of the commercial success of expert systems?*", and (b) "*interrogative + a statement*", e.g., "*Why do different semantic images emerge when a form is sensed by different individuals?*" [46].

7. *Process*. The type of questions asks for a *process* to accomplish something. The interrogatives include "*How…*" and "*What is the way/step/process…*", etc. Most *process* questions follow the following two forms: (a) "*interrogative + a term indicating a process*", e.g., "*What is the process of training a classifier?*" (from the Wikipedia page on AI), and, (b) "*interrogative + a movement*", e.g., "*How does a semantic link network derive a potential semantic link?*" and "*How to predict the evolution of semantic link network?*" [46].

Appendix C. Patterns of answering question

1. Replacement match

An answer matches a question by replacing the interrogative with one or several semantic objects or a statement. For example, for the question "*What do the basic semantic links in the physical space include?*", the answer is "*The basic semantic links in the physical space include the spatial relation, movement relation, symbiosis relation, energy transformation relation and material flow.*" [46].

The pattern has the following characteristics.

(1) It concerns the semantic objects and semantic links surrounding interrogative and ignores other relatively irrelevant part. For example, the sentence "*Artificial intelligence is intelligence displayed by machines, in contrast with the natural intelligence displayed by humans and other animals.*" (from the Wikipedia page on AI) can answer the question "*What is artificial intelligence?*" because the interrogative can be replaced by "*intelligence displayed by machines*". The semantic objects and semantic links in "*in contrast with the natural intelligence displayed by humans and other animals*" are relatively irrelevant to the question.

(2) It allows extension of semantic object through semantic link. For example, for the question "*What are the main problems of artificial intelligence?*", the answer "*Natural language translation is a problem*

of computer science.” does not meet replacement match pattern precisely. But the answer and question still meet the pattern because “*artificial intelligence*” has a semantic link “*subcategory*” to “*computer science*”. It is more understandable if the extension is done through a direct semantic link.

The pattern is usually suitable for the following questions:

(1) question about *object description*;
(2) question about *target object*;
(3) question about *time* “*interrogative + a term indicating a time*”;
(4) question about *location* “*interrogative + a term indicating a location*”;
(5) question about *person* “*interrogative + a term indicating a person*”;
(6) question about *reason* “*interrogative + a term indicating a phenomenon or an event*”; and,
(7) question about *process* “*interrogative + a term indicating a process*”.

2. Common-statement match

A statement at least contains one predicate and one semantic object which is the subject or object of the predicate. An answer matches a question because (a) they share a common statement, (b) the interrogative indicates the complement of the predicate in the statement, and (c) the type of the complement matches the type of the question. For example, the sentence “*In 1935, Turing conceived the modern computer for the first time.*” [46] can be considered as the answer to the question “*When did Turing conceive the modern computer?*” because: (a) the sentence and the question contain a common statem*ent “Turing conceives the modern computer*”, (b) the question asks for *time* of the statement, and (c) the sentence contains “*1877*” which indicates the *time* of the statement. In this pattern, the matching relations between the usage of implement and type of question need to be pre-specified.

The pattern is often used for “*interrogative + an event*” of *time* question, “*interrogative + an event*” of *location* question, “*interrogative + a description or an action of a person*” of *person* question, and *target object* question.

3. Semantic-link match

Some ways to answer questions require certain kinds of semantic links. An example is that people usually use three ways to answer an *object description* question. The first way is to describe a semantic object by subordinating it in a super-category or giving its subcategories. The second way is to list the parts of a semantic object. The third way is to illustrate the instances of a semantic object. The three ways need to know the *subcategory*, *is-part-of* or *instance* links that connect to the semantic objects in question and answer. For example, for the question “*what is artificial intelligence?*”, the three answers from the Wikipedia page on AI respectively correspond to the three ways, which are “*The AI field draws upon computer science, mathematics, psychology, linguistics, philosophy, neuroscience, artificial psychology and many others.*”, “*Artificial intelligence research includes robotics, natural language processing, expert system, etc.*” and “*Some artificial intelligence applications can beat human, such as AlphaGo*”. The three sentences can answer the question because the semantic objects “*computer science*”, “*mathematics*”, “*psychology*”, “*linguistics*”, “*philosophy*”, “*neuroscience*”, “*artificial psychology*”, “*robotics*”, “*natural language processing*”, “*expert system*”, and “*AlphaGo*” have “*subcategory*” link to “*artificial intelligence*”. Besides the three types of semantic link, the way to answer a question may require other kinds of semantic links. For example, for a question that asks for the synonym of a semantic object, an answer must contain a semantic object that has *synonym* link to the semantic object in the question.

Besides *object description* question, some *target object* questions (e.g., the question asking for synonym) can be answered by the pattern. The pattern requires external knowledge to provide the semantic links between semantic objects in question and answer.

4. Sentence-meaning match

In replacement match pattern, common statement match pattern, and semantic-link match pattern, a sentence answers a question by using the partial information of answer. But some sentences can answer a question because their overall meanings meet the question. An example is the fourth way to answer an *object description*

question besides the three ways in semantic-link match pattern, which is to describe the connotation of a semantic object. For example, for the question "*What is artificial intelligence*?", the answer "*Artificial intelligence is to develop a device that perceives its environment and takes actions that maximize its chance of success at some goal.*" describes the connotation of "*artificial intelligence*". Most *reason* questions and answers follow the pattern. Some *process* questions that only need one answer sentence are answered by the pattern as well.

5. Multi-sentence meaning match

An answer matches a question if the answer contains multiple sentences or clauses, and the sentences or clauses have *sequential* relations and the overall meaning of them meets the question. Specifically, an answer matches a question because of satisfying the following four conditions: (a) the question asks for a process to achieve something or the components of something; (b) the answer contains multiple sentences or clauses, (c) the meanings of the sentences or clauses are about time, space, process or logic, and (d) the sentences or clauses meet the requirement of the question as a whole. Most *process* questions are answered by the pattern. Some *reason* questions that need multiple steps of inferring are answered by this pattern as well.

Appendix D. The functions of the semantic link network

1. *Filtering function*

If a semantic link network contains the semantic links between semantic objects within text and some certain types of categories (e.g., *time*, *person*), the function can filter an answer range by removing the sentences that do not have any semantic object whose type matches the type of question. For the 7 types of question, the function can filter most unfit sentences for *time*, *person* and *location* questions. For different patterns, the function is helpful for the 5 patterns as long as a question asks for the answer involving some certain kinds of semantic objects. Filtering function relies on establishing the semantic links between semantic objects in text and some types of categories.

2. *Expanding function*

The function can help identify more semantically-related answers by semantically expanding the meaning of semantic objects in question and answer. For example, for *target object* question "*What do the problems of AI research include*?", the sentence "*The traditional problems of AI research include reasoning, knowledge, planning, learning, natural language processing, perception and the ability to move and manipulate objects.*" (from the Wikipedia page on AI) answers the question by replacement match pattern. By tracing the semantic links between "*artificial intelligence*" and "*natural language processing*", and between "*natural language processing*" and other semantic objects, the sentence "*Challenges in natural language processing frequently involve speech recognition, natural language understanding, and natural language generation.*" can be considered as an answer as well however its relevancy to the question is not as close as the first answer.

Expanding function can help the 7 types of question and the 5 patterns in different degrees. Specifically, the function can help *objects description* and *target object* question because answering the two kinds of questions often requires classification knowledge. The function can help replacement match pattern and common statement match pattern because the two patterns need to determine two semantic objects or two semantic links are equivalent or not and the three types of semantic link can expand the range of equivalent semantic objects and semantic links, and the function can help semantic-link match pattern because the way to answer some questions directly requires the three types of semantic link.

3. *Matching function*

The function can match answer and question when answering a question implies the existence of the three types of semantic link (or a chain of the links) between semantic objects in question and answer. For example, semantic link network can help answer the questions on specialized or generalized concept as people tend to build a concept through its subcategories, super-categories, instances, parts and whole. So the function is helpful for semantic-link match pattern.

As for other patterns, by using the semantic link network in the paper, the function can hardly help sentence meaning match pattern because the pattern requires other types of semantic link besides *subcategory*, *is-part-of* and *instance* link. On one hand, the overall meaning of a sentence often involves other types of semantic links. On the other hand, the way to match question and answer often requires other types of semantic link. For example, the way to match an answer and a *reason* question often requires *inferring* link. The semantic link network can hardly help multi-sentence-meaning match pattern because of the same reasons.

Appendix E. Construction of the semantic link network based on Wikipedia

The semantic link network on an answer range A is constructed by following steps.
1. Get the terms $K = (k_1, k_2 \ldots k_n)$ from A which is elaborated in Appendix F.
2. Get the semantic objects indicated by the terms, denoted as $C_{start} = \{c_1, c_2 \ldots c_m\}$. For each term r in K, if r indicates a page and a category, only one semantic object is added in C_{start}; if r indicates neither a category nor a page, the term is ignored.
3. For each semantic object c in C_{start}, calculate $Up(c) = \{c_i \mid$ if c has a semantic link chain to $c_i\}$ by iteratively upward-searching superior semantic objects starting from c until reaching the top object in SLN_{wiki}. SLN_{wiki} is the semantic link network that covers the semantic objects in Wikipedia.
4. Construct and return a semantic link network $SLN(A)$ in which the nodes are $Up(c_1) \cup \cdots \cup Up(c_m)$ and the semantic links between the semantic objects are extracted from SLN_{wiki}.

Step 3 has the two details as follows.
1. Breaking cycle. It is self-contradictory to have a cycle in the semantic link network. However, the semantic links in Wikipedia sometimes form a cycle. For example, Wikipedia has *cognitive science → cognition → thought → mental processes → cognitive science.* To break the cycles, each visited category is assigned with a label for avoiding being re-added to the constructed semantic link network.
2. Eliminating nonacademic or irrelevant semantic objects. Sometimes Wikipedia categories are too casual for specialized knowledge. For example, *artificial intelligence* is a subcategory of *emerging technologies* whose connotation and extension changes with time. Sometimes Wikipedia connects domain-irrelevant semantic objects. For example, *emerging technology* indirectly connects *artificial intelligence* and *Stem cell research*, which are subcategories of *emerging technology*. To eliminate nonacademic or irrelevant semantic objects, a WordNet-based branching-cutting algorithm is performed.
    While searching the superior semantic objects of a semantic object e, the following steps are performed:
    1. find the superior semantic objects of e in Wikipedia,
    2. if the number of the superior semantic objects is greater than *numberSuper*, calculate the average semantic relatedness between each superior semantic object and the semantic objects in C_{start}, the semantic relatedness measure [43] is provided by WordNet::Similarity (http://wn-similarity.sourceforge.net/),
    3. the superior semantic objects with the first *numberSuper* average semantic relatedness are reserved.

Appendix F. Terms extracting model and Filtering function

For the texts in an answer range, the terms (key terms) are obtained by NLP tool [34], and there are 3 auxiliary ways to get the terms according to text sources: 1) given by the author, e.g., keywords of paper, 2) tagged by reader, e.g., tags of webpage, 3) assigned by third parties, e.g., IEEE online library or ACM online library.

The filtering function first identifies the type of question and then filters answer range. Each sentence within an answer range is assigned a type set which records the types of the appeared semantic objects in the sentence. For example, for a sentence "*The field of AI research was born at a workshop at Dartmouth College in 1956.*" (from the Wikipedia page on AI), its type set is {*location, time*}. So the sentence can answer some *time* questions and *location* questions. The type set of a sentence does not include *object description*, *target object*, *reason* and *process* because filtering function can hardly help answer the 4 types of question. So filtering function does not occur when a question belongs to the 4 types. In the experimental QA system, filtering function is implemented with the Named Entity Recognition and certain categories (time, location, etc.) in Wikipedia.