

Typology of Ambiguity on Representation of Information Needs

The first part of the paper develops a framework explaining the need to disambiguate user inquiries to improve information systems and services. Theoretical grounds for this framework are explained for how questions are categorized on the basis of their ambiguity type, while the relevant literature is reviewed including both the traditional and the digital information service environment. The second part of the paper categorizes a set of questions (400 Qs), originally collected for TREC 8 and 9 QA Tracks, according to ambiguity type. Three types and two dimensions of ambiguity are identified by the author with the acceptable levels of inter-coder agreements presented. The last part of the paper discusses three aspects of information systems and services, mainly related to user-system and user-information intermediary (i.e., a reference librarian) interactions, on the basis of the results of categorization. Those three aspects include (1) increasing user input to make initial queries less ambiguous, (2) reducing search space by disambiguating queries, and (3) clustering search results based on the characteristics of prospective answers. In each of the three aspects, discussions on the evolving environments of virtual reference services were presented.¹

Disambiguating human inquiries, either in a semantic or lexical approach, is an essential process to consider in developing information systems and services. This paper discusses this process for design in two related domains—information systems and services—but in a specific aspect of such domains—accommodating different types of full-sentence questions.

The information system domain attempts to refine question categorization to develop question-answering (QA) systems. While significant work has been done in this area, consideration of question ambiguity has been limited on classifying questions. This paper presents a classification of a set of full-sentence questions originally collected for the Text REtrieval Conference (TREC) 8 and 9 Question Answering (QA) Tracks, according to their ambiguity which could mislead an engaged information system.² The information service domain concerns situations in which prospective users are engaged in the searching activity with the information needs represented in the question set. The discussion then extends into the possible intervention of a human information intermediary (i.e., a reference librarian) in the searching process.

On the basis of the types and

Yang-woo Kim

Yang-woo Kim is Assistant Professor, Division of Knowledge and Information Studies, Hansung University, Seoul, Korea.

This paper is a fully revised version of an earlier work presented at the Annual Conference of the American Society for Information Science and Technology, Long Beach, California, October 19–22, 2003.

Reference & User Services Quarterly, vol. 53, no. 4, pp. 313–25
© 2014 American Library Association. All rights reserved.
Permission granted to reproduce for nonprofit, educational use.

dimensions of ambiguity identified, three aspects of information systems and services are discussed mainly related to *user-system* and *user-information intermediary interactions*. Those three aspects are (1) increasing user input to make initial queries less ambiguous, (2) reducing search space by disambiguating queries, and (3) clustering search results on the basis of characteristics of prospective answers.

Unlike the majority of question analyses conducted on the previous work (reviewed in this paper), this study does not aim to categorize questions according to plausible inference, anticipating a single answer to a question. Instead, users' query statements are classified *on the basis of what the author did not explicitly know about the inquirers' intentions*. This approach seems reasonable because what is manifestly known of an inquirer's intention from a single sentence query is quite limited. In addition, the increase in fact-finding questions in the digital environment provides significance for this specific study while the relevant literature indicates an increase in the virtual reference questions compared to the decrease in traditional reference questions.³

This paper, therefore, addresses the following research questions:

- What are the different types of ambiguity in a set of questions, originally collected for TREC 8 and 9 QA Tracks?
- What are the implications of the ambiguities identified for user-system and user-information intermediary (i.e., a reference librarian) interactions?

BACKGROUND

Researchers have attempted to categorize questions (or user needs) with varying approaches from related fields. The review of relevant literature indicates little consideration of sentence ambiguity, particularly in categorizing an exhaustive set of questions.

Internal Need vs. Expressed Need

Many studies discussed possible discrepancies between people's internal needs and expressed needs. Taylor suggested the need to accommodate *the users' hidden needs*; he presented four different types of user needs as levels of question formation: visceral, conscious, formalized, and compromised.⁴ Several authors further developed Taylor's ideas, emphasizing the need to cope with the discrepancies between the internal (visceral, conscious, formalized) and the expressed (compromised) needs. Ingwersen emphasized the importance of identifying the relation between the formalized need and the compromised need. The compromised need (the question as presented to librarian or system) is an expressed need. When there are discrepancies between the internal and the expressed needs, there seem to be stronger possibilities of ambiguity in users' questions.⁵ In a similar sense, Stevens indicated the need to determine whether users are seeking

more or less than what they have requested.⁶ In particular, he indicated that the question a user asks may not be the question he/she wished to be answered. This is certainly a situation in which a reference librarian or a system designer needs to cope with the ambiguity of the user question.

Categorizing Questions: Nonsystem Approaches

Substantial work can be found on question categorization that is not directly associated with a specific system use or evaluation. Several authors have discussed the *underlying meaning of questions* with a limited number of arbitrary examples and without an extensive question set.⁷ These discussions included inferring the *inquiry scope* of the questioner—Belnap and Steel, Harrah; and underlying meanings of questions with possible responses—Graesser and Black.⁸ In particular, Belnap and Steel emphasized the importance of inferring the questioner's intention when it is not very clear, in other words, ambiguous. Meanwhile, Graesser and Black addressed internal meanings of questions when there can be vague meanings. Yet, as indicated, these studies did not extend their research scope into an extensive question set, illustrating a few arbitrary examples. Categorizations of questions in these studies were primarily based on different types of interrogatives.

In the traditional library use environment, a significant body of work has been conducted on classifying *reference questions*. Categorizations of exemplary approaches are based on the following: (1) answer format that satisfies users—Heiber; (2) types of need (i.e., direction, information, and general reference)—Seng; (3) presupposed concepts—Derr; (4) type of sources requested (i.e., general or specific)—Brown; (5) complexity of questions—Robinson; and (6) a taxonomy of research questions related to information literacy—Cordell and Fisher.⁹

Although these studies accommodated a variety of question types, they did not examine the questions' ambiguity types.

Categorizing Questions: Digital Reference Environment

More recently, another body of work discussed classification of reference questions in the emerging environment of the digital reference services. Categorizations of exemplary approaches are based on the following: (1) the types (i.e., patrons' rudeness or abusiveness, poor writing skills) of problems in providing virtual reference service—Lindbloom et al.; (2) types (i.e., directional, ready reference, specific search and research, policy and procedural or holdings/do you own) of synchronous virtual reference questions at a university library—Arendt and Graves; (3) the question types (i.e., technical problem, directional/policy, known item, facts/ready reference) in virtual environment—Marsteller and Mizzy, De Groote, Fennwald; and (4) types of information sought (i.e., topic, background, search history, search history, extent-depth) by librarians in live chat virtual reference—Radford et al.¹⁰

As discussed, while the relevant literature indicates an increase in the digital reference questions compared to the decrease in traditional reference questions, the increase in short factual questions in the digital environment provides significance for this specific study. This is because what is manifestly known of a user's intention from a short factual query is quite limited. In other words, there can be more ambiguous questions in such circumstances. Nevertheless, the above studies did not address their ambiguity types, although they provided a variety questions and problem types. This could be the shortcomings of the above studies. This strengthens the importance of this specific study.

THEORETICAL PERSPECTIVE

While utilizing the interrogatives of questions can be an effective tool for eliciting an inquirer's needs, it is inadequate as the sole component of such elicitation. This study attempts to explain this inadequacy in relation to an essential limitation of current information systems, mainly originating from the term occurrence-based retrieval mechanism. One typical drawback of such a mechanism is to yield a low *precision* result at the cost of high *recall*.

As information retrieval (IR) literature indicates, the term *occurrence-based retrieval mechanism of full-text searching*, in which systems are able to search for the occurrence of any single word or phrase in full-text documents, yields a high number of irrelevant returns for each search, burdening end users with sorting out an enormous set of items. In many of these instances, a document is returned solely because it has a search term (usually a topical query term) somewhere within the document but possibly within a completely different context from what the user actually needed. This limitation remains when a system adopts a matching technique, a more traditional method of IR. An information system that relies heavily on topic-based index terms is more effective in promoting recall than improving precision, while such a reliance is a common shortcoming on a majority of computerized bibliographic databases.¹¹

While ongoing efforts in the IR community have attempted to improve this drawback (low *precision* at the cost of high *recall*), the progress seems only moderate. Current systems have a very limited capability to understand the analytic aspects of texts, i.e., the underlying meanings of texts in either query or document side, although some progress has been made in identifying such text characteristics based on linguistic features.

Figure 1 depicts such a limitation. An underlying assumption of this model is that an information system can recognize only a limited portion of what users think, labeled as the Represented Zone (RZ) on the left side of the figure. Here, it is illustrated as the visible portion of an iceberg above the water surface. The Unrepresented Zone (UZ) below the water surface, which accommodates the underlying meaning of user need, cannot be recognized by the system, as it is left invisible.

The limited space of the RZ at top right accounts for a confined portion *utilized* during the retrieval process, out of the entire VALUE of the *system* or *resource* stored within the system. This is explained as follows. On the left side of the figure, a limited portion of user need was represented and input either because of user's limited skill or insufficient system features; thus the unrepresented need was not reflected in the retrieval process. On the right side of the figure, the potential value of the *resource* stored in the system was utilized only to a limited extent because of the shortcoming of the retrieval method that relied on the term occurrence. Accordingly, the invisible portion on the right side of the figure accounts for an unutilized value of system or resource stored on the system.

An important goal of interactive IR systems should be to devise mechanisms to connect the two areas at the bottom, Unrepeated and Unutilized zones on the left and right sides of the figure, respectively. This paper attempts to gain ideas to improve these mechanisms with respect to user-system interactions and the intervention of an information intermediary (i.e., a reference librarian) on such interactions. More detailed components of such interactions discussed include representations (both in the user- and the system-side with an emphasis on the former), interface designs, retrieval mechanisms, and intermediary interventions. In a related study, Kim examined a broader domain of the four different zones with an extended iceberg model presented.¹²

THE QUESTION SET OF THE STUDY

The TREC 8 and 9 QA Track Question Set

For this study, two question sets of the TREC 8 and 9 QA tracks were used as the original sources of data. To have the test set represent a wide spectrum of subjects and question types, the TREC question sets were collected from four different sources: TREC QA participants, the National Institute of Standards and Technology (NIST) TREC team, the NIST assessors, and questing logs from the FAQFinder system.¹³

When these question sets were created, any questions that were judged as ambiguous were eliminated to create a *set of clean, straightforward questions and answers*.¹⁴ Nevertheless, the results of the system performance showed that the questions had more different answers than anticipated.¹⁵ So, it was meaningful to examine the QA track questions with respect to their ambiguity types.

The Selected Question Set of This Study

For this study, 400 questions were used from the original question sets, 200 of which were from TREC 8 QA Track and another 200 from the selection of 693 questions from TREC 9 QA Track. Those 200 questions chosen as the latter were selected in numerical order. The full list of 400 questions is available on the TREC homepage.¹⁶

It is believed that the question set of 400 queries has a reasonable level of exhaustiveness for an analysis. In a library

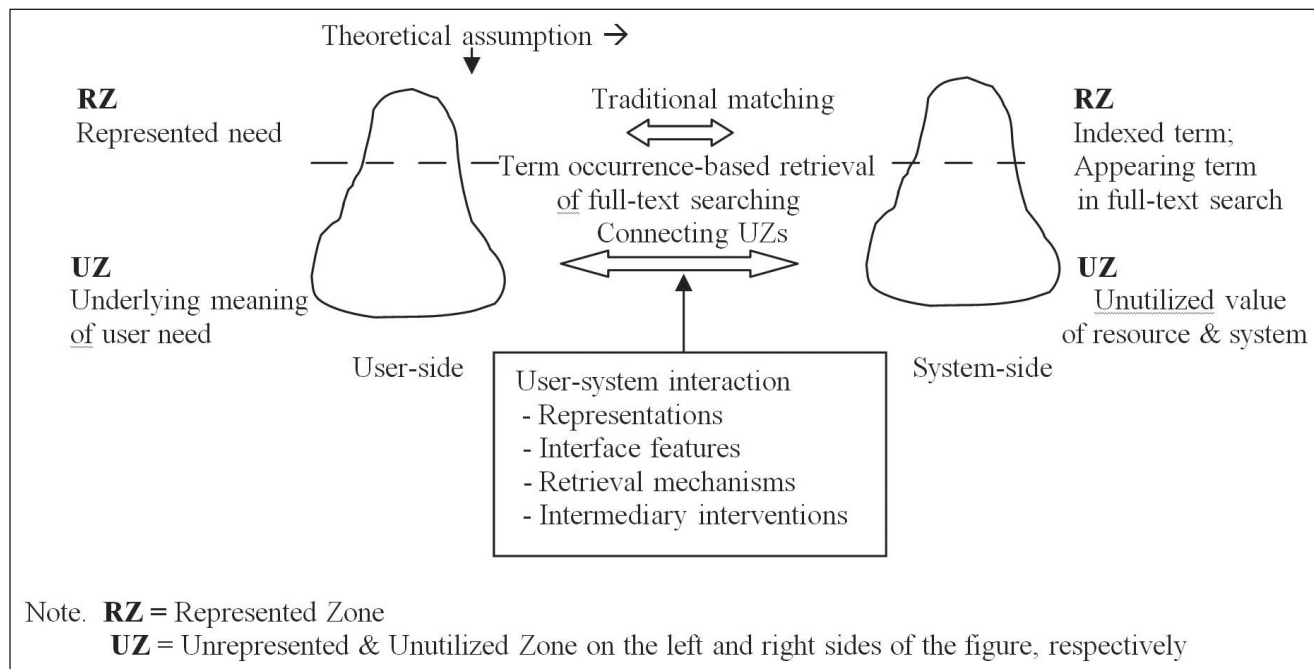


Figure 1. An iceberg model of Interactive IR to accommodate the underlying meaning of user needs: building a bridge over Unrepresented and Unutilized Zones (UZs)

setting, for instance, the questions could be the initial inputs of 400 patrons for user-librarian interactions.

MAJOR CONCEPTS ON CATEGORIZATION

This section explains major concepts adopted by the author in the creation of question categories. A few exemplary questions from the question sets are presented for each concept. See the Results section for further details.

Types of Ambiguity

Three types of ambiguity were presented on categorization. *Semantic ambiguity* typically occurs when one is faced with vagueness about adopting a *meaning of a word or a sense of a single meaning*. The origins of ambiguity identified in this study for this type include Polysemy, the different related senses of a single meaning, and Homonymy, literally different meanings. An exemplary question for Polysemy is “How rich is Bill Gates?” For questions such as this, there can be a few possible answers namely, *the person’s total wealth or his ranking among the top rich in the world*. In addition, an exemplary question for Homonymy is “Who leads the star ship Enterprise in Star Trek?” For such a question, there are a few alternative answers such as *a particular character in the movie and an actor chosen for the role* (see table 1 and appendix A for further examples).

Syntactic ambiguity differs from semantic ambiguity in that no term in a sentence has more than one meaning or different

senses of a meaning. Instead, the ambiguity originates from a vague sentence structure, requesting one to select between alternative grammatical structures. A typical pattern in this type of ambiguity is associated with lack of clarity in determining modified terms. An exemplary question in this category is “what is the name of the rare neurological disease with symptoms such as involuntary movements (tics), swearing, and incoherent vocalizations (grunts, shouts, etc.)?” Depending on how the terms are modified by the phrase “such as” (i.e., either one or all symptoms), the possible answers can vary.

Pragmatic ambiguity typically occurs when not all (none) of the propositional content of a sentence is explicitly given. This type of ambiguity is not attributable to different meanings of words or grammatical structure, but rather to how things exist in the real world. An exemplary question of “what is the brightest star visible from Earth?” shows this type of ambiguity. Here, the term *visible* can also generate two different contexts: *visible when an instrument is used or visible to the unaided eye*. In this example, the meanings or grammatical structure do not influence the sentence ambiguity. The lack of specification on the conditions of human perceptions—with or without use of instrument—causes this ambiguity, while essential terms such as *brightest* or *visible* have one single identical meaning.

The Dimensions of Ambiguity

Two different dimensions of ambiguity are presented below as subcategories of the categorization.

Context-based ambiguity occurs when one can think of

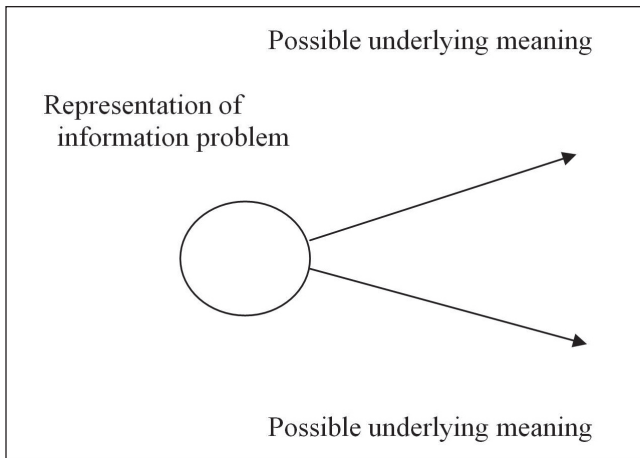


Figure 2. Context-based ambiguity with more than one principal focus

alternatives as underlying meanings to a question that are not necessarily related to each other. In other words, this ambiguity occurs when one can plausibly suggest more than one possibility of meaning for different situations or circumstances (see figure 2).

The question “who is the leader of India?” belongs to this dimension because the two possible alternatives (in appendix A shown as *the President* and *the Prime Minister (PM) of the country*) involve two different individuals. The president is likely to play a symbolic role whereas the PM has the genuine political power in the cabinet system of the country.

In addition, *dynamic* and *static contexts* have been adopted to further characterize this dimension of ambiguity.¹⁷ *Static* context refers to a situation that is fixed over time, situations, and environment, while *dynamic* context is changeable.¹⁸ For example, the three possible alternatives for the question, “How many Vietnamese were there in the Soviet Union?” illustrate these two contexts. While the corresponding figures in the beginning and at the last moment of the regime can be labeled as *static*, figures from the whole period is certainly changeable, demonstrating its *dynamic* nature (see appendix A). The discussion section of the paper further explains the searching process for information needs in these different contexts.

Scope-based ambiguity concerns vagueness to the extent that a question should be expanded or narrowed from a principal inquiry focus; that is, vagueness in the intended range of inquiry domain (see figure 3).

The question “when did communist control end in Hungary?” was categorized into this dimension. One plausible interpretation is that the inquirer wants to identify a period of the particular era as a primary interest of the inquiry. This could be considered the principal focus of the inquiry. The factor that characterizes the end of the communist regime (i.e., the collapse of Soviet Union) will be a possible alternative extension of the principal focus. Here, the extended focus relates to the principal focus (see table 1 and appendix

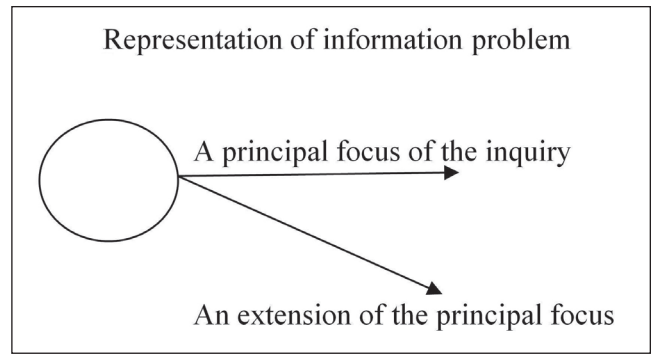


Figure 3. Scope-based ambiguity with one principal focus

A). The question “when did the Jurassic period end?” was categorized into this dimension as well. Again, a plausible interpretation is that the inquirer wants to identify a period of the particular era as a primary interest of the inquiry. In fact, the TREC QA track selected a specific time (130 million years ago) as the correct answer to this question. This could be considered the principal focus of the inquiry. The factor that characterizes the end of the Jurassic period (i.e., disappearance of dinosaurs) will be a possible alternative extension of the principal focus. Here, the extended focus relates to the principal focus. In general, the characteristics of a historic era should have a close relationship with its chronological period (see appendix A).

Types of Specificity on Regional and Time Space

The concept of *regional and time space* were used as subcategories for *the scope dimension of pragmatic ambiguity*. In each space, two types of specificity (vertical and horizontal) were presented to indicate the origin of the ambiguity. The ambiguity here originates from vagueness of the specificity in answering an inquirer’s question.

The vertical specificity on regional space concerns the degree that specifies the single space of an occasion (or an entity) already mentioned in the question. For example, the question “where did the 6th annual meeting of Indonesia-Malaysia forest experts take place?” belongs to this category. The vertical continuum in figure 4.1 illustrates the different degrees of specificity for a single space.

The horizontal specificity on regional space concerns the selection of a single space among multiple possibilities, as the regional space has not been clearly stated in the question. The question “what is considered the costliest disaster the insurance industry has ever faced?” belongs to this category (see figure 4.2 for an example of the horizontal continuum).

The idea for *vertical specificity on time space* is basically the same as regional space. Here, the question has vagueness to an optimal degree to specify a single space. The question of “When was General Manuel Noriega ousted as the leader of Panama and turned over to US authorities?” belongs to this category (refer to figure 5.1 for an example of vertical

FEATURE

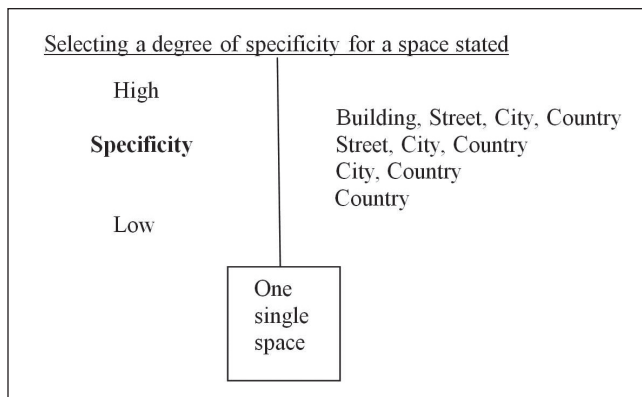


Figure 4.1. Specificity of regional space: an example of vertical continuum

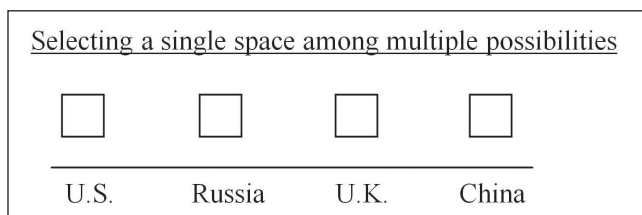


Figure 4.2. Specificity of regional space: an example of horizontal continuum

continuum).

Similar to regional space, *the horizontal specificity on time space* deals with the designation of a single space among many different spaces. For instance, the answer to the question “What was the name of the US helicopter pilot shot down over North Korea?” can reflect more than one time space in a horizontal continuum (see figure 5.2).

RESULTS

This section presents categories of questions identified as ambiguous from the selected question sets. A total of 80 questions were labeled ambiguous by the author from among 400 questions, with a variety of ambiguity types.

The *coefficients of reliability* were computed to measure *inter-coder agreement* between the author and two independent coders recruited.¹⁹ The formula used is $C.R. = 2M / (N1 + N2)$, where M is the number of coding decisions on which the author and a coder are in agreement, and N1 and N2 refer to the number of coding decisions made by them. The reliability measures reached an “acceptable” level with the ratio of .82 and .85, respectively.²⁰

The results suggest the need for careful consideration when accommodating the vaguely expressed intentions of end users. Table 1 presents the categories of ambiguity with an exemplary question shown for each category. The total number of ambiguous questions was 87 because 7 questions

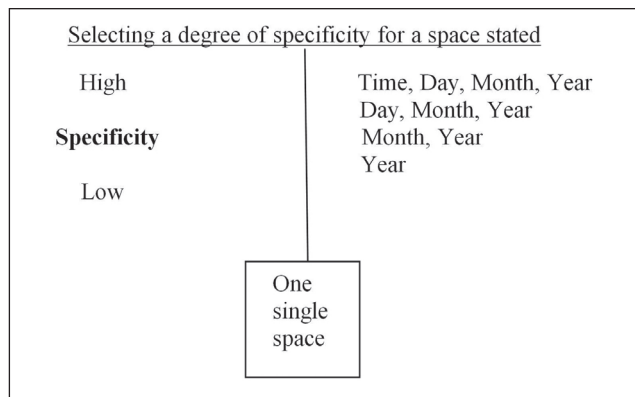


Figure 5.1. Specificity of time space: an example of vertical continuum

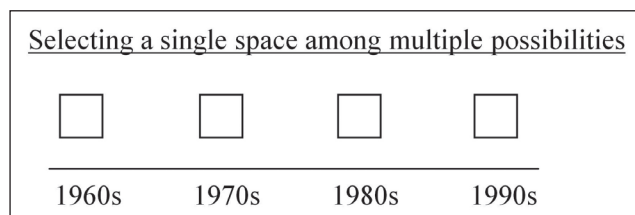


Figure 5.2. Specificity of time space: an example of horizontal continuum

actually belonged to more than one category and were thus counted twice (indicated by parentheses). The *focus of plausible alternatives* illustrates possible underlying meanings for each question.

Appendix A addresses further details with all questions belonging to each category presented. Here, the characteristics of prospective answers are presented in a separate column, further clarifying the focus of plausible alternatives. In addition, characteristics of context indicate the static and dynamic nature of the question for context-based ambiguity.

DISCUSSION

These results suggest further progress is necessary on the following three aspects of information systems and services to enhance *user-system interaction* and *user-information intermediary* (i.e., a *reference librarian*) interaction. While the former (user-system interaction) mainly deals with improving currently available information systems, including their capacity to provide appropriate interface for natural language query inputs, the latter (user-information intermediary interaction) concerns the use of current systems, facilitating the searching process of the end user.

Increasing User Inputs

The first aspect concerns increasing user-inputs to *make initial*

Table 1. Typology of Ambiguity on Representation of Information Needs

Types of Ambiguity	Dimension of Ambiguity	Origin of Ambiguity	Nr. Of Q	Example of Question	Focus of Plausible Alternatives
Semantic			19 (7)		
	Context	Polysemy—degree	2 (1)	How rich is Bill Gates?	Total wealth; ranking among the top rich in the world
		Polysemy—quantity/size	4	What country is the world's leading supplier of cannabis?	Amount; value, etc.
		Polysemy—monetary term	1	What debts did Quintex group leave?	Financial liability; nonfinancial liability
		Homonymy—acting body	5	Who leads the star ship Enterprise in Star Trek?	Character in the movie; actor chosen for the role
		Homonymy—entity	1 (1)	Where is the Taj Mahal?	Monument; hotel; restaurant
		Homonymy—Location	1	Where did Bill Gates go to college?	Name of college; name of city
	Scope	Polysemy—occurrence	5 (5)	When did communist control end in Hungary?	Time period (Date)—circumstance; characteristics; occasion
Syntactic			2		
	Context	Multiple modified terms	1	What costume designer decided that Michel Jackson should only wear one glove?	What-who; what-field
	Scope	Extent of modified examples	1	What is the name of the rare neurological disease with symptoms such as involuntary movements (tics), swearing, and incoherent vocalizations (grunts, shouts, etc.)?	Fulfilling all examples (i.e., swearing) listed; fulfilling part of them
Pragmatic			66		
	Context	Condition of human perception	1	What is the brightest star visible from Earth?	Aided eye; unaided eye
		Time space of inquiry	3	How many Vietnamese were there in the Soviet Union?	In the beginning; in the last stage; during the whole period
		Domain of regulation	1	What is the legal blood alcohol limit for the state of California?	Transportation tools; sports activity
	Scope	Regional space—VS *	17	Where is it planned to berth the merchant ship, Lane Victory, which Merchant Marine veterans are converting into a floating museum?	Country—city; street; building, etc.
		Regional space—HS **	4	What is considered the costliest disaster the insurance industry has ever faced?	US only; any countries
		Time space—VS	22	When was General Manuel Noriega ousted as the leader of Panama and turned over to US authorities?	Decade—year—month; day—time
		Time space—HS	2	What was the name of the US helicopter pilot shot down over North Korea?	One particular helicopter; all helicopters
		Time space—extent of specified space	4	How much could you rent a Volkswagen bug for in 1966?	Amount as of the specified year; value converted as of current year
		Level of details—Terminology	9	What is Head Start?	Simple definition; detailed explanation
		Level of details—Reason	1	Why did David Koresh ask the FBI for a word processor?	A simple reason; detailed explanations
		Level of details—Product	1	What does the Peugeot company manufacture?	A major product; line of products
Total			87 (7)		

SOURCE—Original sources of questions: TREC 8 and 9 QA Tracks.

NOTE—Number of questions that belong to more than one category is in parentheses.

VS * = Vertical Specificity. HS ** = Horizontal Specificity.

query inputs less ambiguous. A key concern here will be to induce users to present more information about their information need, while facilitating them in representing their needs.

For *user-system interaction*, the results suggest the need to improve the interface features of information systems in the following aspects, while retrieval mechanisms should be improved to support such features: (1) leading users to specify a wider range of selection criteria in the query formulation as many commercial web search engines do, yet in a limited extent, and (2) providing users with an increased space for query input as many virtual reference bulletins do. Improvements on such aspects will help systems clarify users' needs.

For *user-information intermediary (i.e., a reference librarian) interaction*, the intermediary's use of *open questions* in an early stage of face-to-face interaction is a possible technique dealing with drawbacks (limited representations of user needs). Such limitations or "compromising," in Taylor's term, are affected by various components: (1) limited (searching) skills of users, (2) misperceptions of system capacity and functions by end users, and (3) actual shortcomings of system features.²¹ The appropriate intervention of the intermediary is necessary to deal with each component. In the virtual reference interaction, the intermediary's intervention can be limited because of little visual and auditory cues—e.g., the user's age group, gender, ethnic background, and physical location. However, a prearranged guideline from a library site can possibly instruct users to input their initial questions more clearly.

Reducing Search Space by Disambiguating Queries

The second aspect deals with refining the inquiry scope (reducing search space in the system-side) by disambiguating user needs already represented.

With respect to *user-system interaction*, enhancing the interactive dialogue between the two sides is a prospective area of further research and development. One possibility begins with identifying meaningful (language) characteristics of query statements at an initial stance before the system searches through its full-text documents. Such identification is a prerequisite for an information system to effectively *ask back* to its end user about vague query input. For example, with the input of a question such as "what is the brightest star visible from Earth?" a system could yield an improved return if it is supported by a mechanism that directly asks back "which of the following did you mean: to naked human eye or to scale?"

Another possibility utilizes the meaningful characteristics of prospective answers located after a system searches through full-text documents to further identify the query's ambiguity factor. The identification of such characteristics as *unit of tons or dollars* can be an example for the question "what country is the world's leading supplier of cannabis?" When an information system is supported by a mechanism for such identification, the next necessary step would be to extend it for an interaction with a system user. For the above question, a system can better understand the user's need by

asking back "do you mean by *amount* or *value*?" Appendix A shows additional examples.

The above two examples suggest that the ambiguity types can yield fruitful schemes for classifying questions in the process of answering questions automatically.

Concerning *user-information intermediary (i.e., a reference librarian) interaction*, the intervention of the intermediary in traditional reference setting would be more effective when accompanied by the use of specific closed questions (i.e., "Do you mean by *amount* or *value*?"). Again, the intervention can take place either at an initial interaction with a user or at a later stage of interaction after the intermediary has examined the search results. As discussed, the intermediary's intervention is rather limited in virtual reference desk without the sensory cues of face-to-face encounter. In particular, the intermediary can provide only an asynchronous communication unless synchronous chat reference software is used. Perhaps, under the circumstance of asynchronous communication, it would be an appropriate step for an intermediary to present all plausible answers when s/he deals with a short factual question.

Another facet of discussion relates to the contexts (*static* or *dynamic*) of information need represented. When a user's need is identified as a *static context*, the specification of the publishing date for an information resource will be less meaningful, as long as a resource can be located; however, it becomes more significant in a *dynamic context*. The intermediary's immediate recognition of such an aspect of information need should strengthen the interaction with a user, thereby enhancing the searching process (see appendix A for corresponding questions).

Clustering Search Result on the Basis of Answer Characteristics

The third aspect concerns the use of characteristics of prospective answers on clustering search results, attempting again to reduce the search space. Question attributes that have more than one alternative answer indicate the need to refine the clustering process. For instance, the question on the *cannabis supplier* suggests an idea of clustering retrieved items on the basis of such categories as *amount* and *value*.

The intervention of a human information intermediary (i.e., a reference librarian) in the process of a user's interaction with search results should be affected by the availability of clustering features in the searching systems. When an information system has a limited feature for such displays, an intermediary in the face-to-face interaction can directly help the user sort out preferred and plausible alternative answers. In the virtual reference setting, the intermediary can communicate with the end user for such intervention through email, the Internet chat, and other web form-based reference tools such as online reference bulletins.

CONCLUSIONS

To enhance the three aspects of information systems discussed above, further progress needs to be made in the computer-aided techniques of text analysis. In particular, QA systems can automate some answering of reference questions, thus enabling a digital reference service to scale up to handle an increasingly large number of questions. This suggests that digital reference services provide a useful test-bed for future QA systems that implement increasingly sophisticated functionality.²² The extent of improvement of the above aspects will determine the degree or the kind of intervention necessary from the human information intermediary (i.e., reference librarian). For example, user interaction with systems can be affected by three different components and each component can influence the intermediary intervention (as discussed with figure 1): (1) limited searching skills of users, (2) misperceptions of system functions by users, (3) actual shortcomings of system features. In case 1 and 2, the intermediary's (i.e., a reference librarian) direct intervention to cope with ambiguous representation of user needs would be necessary. For 3, the intermediary's knowledge on the inadequacy of the systems would be required to recommend a different system. Overall, the intermediary, as a reference librarian, needs to understand the characteristics of users and systems associated with ambiguous representation of user needs as well as system capacity to deal with such ambiguity.

One possible direction for future study originates from the restrictions on the question set used in this study. Since the question set encompassed only *simple factual questions that had seemingly straightforward answers*, the next step seems to extend the question set to include more complex, analytical questions such as subject-based research type questions. An earlier work in a specific setting (chat reference service) revealed little difference in the effectiveness of question answering between subject-based research and simple factual questions.²³ Yet it will be still meaningful to examine this issue with respect to the question ambiguity to further clarify the impact of such question characteristics.

References and Notes

1. This research was financially supported by Hansung University.
2. The Text Retrieval Conference (TREC) is cosponsored by the National Institute of Standards and Technology (NIST), an agency of the US Commerce Department, and US Department of Defense. TREC consists of several tracks including the QA track. For each TREC QA track, NIST provides a test set of questions. Participants run their own retrieval systems on the data and return to NIST a list of the retrieved top-ranked answers. The QA track is designed to take a step closer to information retrieval rather than document retrieval. The TREC QA track question sets consist of short factual or definitional questions. Related discussions are presented at later sections of this paper.
3. Piritta Numminen and Pertti Vakkari, "Question Types in Public Libraries' Digital Reference Service in Finland: Comparing 1999 and 2006," *Journal of the American Society for Information Science & Technology* 60, no. 6 (2009): 1249–57; Sandra L. De Groot, "Questions Asked at the Virtual and Physical Health Sciences

Reference Desk: How Do They Compare and What Do They Tell Us?" *Medical Reference Services Quarterly* 24, no. 2 (Summer 2005): 11–23.

4. Robert S. Taylor, "Questions-Negotiation and Information Seeking in Libraries," *College & Research Libraries* 29 (1968): 178–94.
5. Peter Ingwersen, "Search Procedures in the Library—Analyzed for the Cognitive Point of View," *Journal of Documentation* 38, no. 3 (1982): 165–91; Norman D. Stevens, "The Importance of the Verb in the Reference Questions," *Reference Librarian* 22 (1988): 241–44.
6. Stevens, "The Importance of the Verb in Reference Questions," 243.
7. Nuel D. Belnap and Thomas B. Steel, *The Logic of Questions and Answers* (London: Yale University Press, 1976); Arthur C. Graesser and John B. Black, *The Psychology of Questions* (Hilldale, NJ: Lawrence Erlbaum Associates, Publishers, 1985); David Harrah, *Communication: A Logical Model* (Cambridge, MA: MIT Press, 1963); Nicholas Rescher, *Empirical Inquiry* (Totowa, NJ: Rowman and Littlefield, 1982).
8. Belnap and Steel, *The Logic of Questions and Answers*; Harrah, *Communication*; Graesser and Black, *The Psychology of Questions*.
9. Caroline Heiber, *An Analysis of Questions and Answers in Libraries* (Lehigh University, 1966); Mary Seng, "Reference Service Upgraded, Using Patron's Reference Questions," *Special Libraries* 69, no. 1 (1978): 21–28; Richard L. Derr, "Questions: Definitions, Structure, and Classification," *RQ* 24, no. 2 (Winter 1984): 186–90; Diane M. Brown, "Telephone Reference Questions: A Characterization by Subject, Answer Format, and Level of Complexity," *RQ* 24, no. 3 (Spring 1985): 290–303; Barbara M. Robinson, "Reference Services: A Model of Question Handling," *RQ* 29, no. 1 (Fall 1989): 48–61; Rosanne M. Cordell and Linda Fisher, "Reference Questions as an Authentic Assessment of Information Literacy," *Reference Services Review* 38, no. 3 (2010): 474–81.
10. Mary-Carol Lindbloom et al., "Virtual Reference: A Reference Question Is a Reference Question Or Is Virtual Reference a New Reality? New Career Opportunities for Librarians," *Reference Librarian* 45, no. 93 (2006): 3–22; Julie Arendt and Stephanie Graves, "Virtual Question Changes: Reference in Evolving Environments," *Reference Services Review* 39, no. 2 (2011): 187–205; Matthew R. Marsteller and Dianne Mizzy, "Exploring the Synchronous Digital Reference Interaction for Query Types, Question Negotiation, and Patron Response," *Internet Reference Services Quarterly* 8, no. 1/2 (2003): 149–66; De Groot, "Questions Asked," 11; Joseph Fennewald, "Same Questions, Different Venue: An Analysis of In-Person and Online Questions," *Reference Librarian* 46, no. 95/96 (2006): 20–35; Marie L. Radford et al., "Are We Getting Warmer? Query Clarification in Live Chat Virtual Reference," *Reference & User Services Quarterly* 50, no. 3 (2011): 259–79.
11. Rebecca Green, "The Role of Relational Structures in Indexing for the Humanities," *Knowledge Organization* 24, no. 2 (1997): 72–83.
12. Yang-woo Kim, "Interactive Information Retrieval Models: Tradition and Development," *Journal of the Korean Society for Information Management* 24, no. 2 (2007): 45–69.
13. Ellen M. Voorhees and Dawn M. Tice, "The TREC-8 Question Answering Track Evaluation," *Proceedings of TREC 8* (2000), accessed April 2, 2013, http://trec.nist.gov/pubs/trec8/t8_proceedings.html; Ellen M. Voorhees, "The TREC-9 Overview of the TREC 9 Question Answering Track," *Proceedings of TREC 9* (2001), accessed April 2, 2013, http://trec.nist.gov/pubs/trec9/t9_proceedings.html.
14. *Ibid.*
15. *Ibid.*
16. The lists of questions from TREC 8 and TREC 9 are available at http://trec.nist.gov/data/qa/T8_QAdata/topics.qa_questions.txt and [http://trec.nist.gov/data/qa/T9_QAdata/qa_questions_201–893](http://trec.nist.gov/data/qa/T9_QAdata/qa_questions_201-893), respectively.

FEATURE

17. M. O'Donnell, "Context in Dynamic Modeling," in *Text and Context in Functional Linguistics*, ed. Mohsen Ghadessy (Philadelphia: John Benjamins Publishing Company, 1999), 63–99.
18. *Ibid.*
19. Ole R. Holsti, *Content Analysis for the Social Sciences and Humanities* (Reading, MA: Addison-Wesley, 1969).
20. Klaus Krippendorff, *Content Analysis: An Introduction to Its methodology* (Beverly Hills, CA: Sage, 1980).
21. Taylor, "Questions-Negotiation and Information Seeking in Libraries," 185.
22. Jeffrey Pomerantz, "A Conceptual Framework and Open Research Questions for Chat-Based Reference Service," *Journal of the American Society for Information Science & Technology* 56, no. 12 (2005): 1288–1302.
23. Nahyun Kwon, "Public Library Patrons' Use of Collaborative Chat Reference Service: The Effectiveness of Question Answering by Question Type," *Library & Information Science Research* 29, no. 1 (2007): 70–91.

APPENDIX A. TYPOLOGY OF AMBIGUITY ON REPRESENTATION OF INFORMATION NEEDS:

All Corresponding Questions Listed

I. Semantic ambiguity

Context-Based Ambiguity

Origin of Ambiguity Questions	Focus of Plausible Alternatives	(Characteristics of) Prospective Answers	Characteristics of Context
Polysemy—monetary term What debts did Qintex group leave?	<ul style="list-style-type: none"> • Financial liability • Non-financial liability (i.e., moral debt; social, economic impact on the industrial sector or business community) 	<ul style="list-style-type: none"> • Figure; Measure units; Monetary amount/value. • Description; Related terms 	<ul style="list-style-type: none"> • Static • Dynamic
Polysemy—quantity/size What country is the world's leading supplier of cannabis?; What country is the biggest producer of tungsten?	<ul style="list-style-type: none"> • Production volume • Sales volume (business profit) 	<ul style="list-style-type: none"> • Measure unit (tons) • Measure unit (money); related term (i.e., market share) 	<ul style="list-style-type: none"> • Both dynamic
Polysemy—quantity/size What company is the largest Japanese ship builder?	<ul style="list-style-type: none"> • Production volume • Sales volume (business profit) • Factory size 	<ul style="list-style-type: none"> • Measure unit (tons) • Measure unit (money) • Measure unit (dock) 	<ul style="list-style-type: none"> • All dynamic
Polysemy—quantity/size What is the largest city in Germany?	<ul style="list-style-type: none"> • Population • Area 	<ul style="list-style-type: none"> • Figure • Different measure unit (i.e., square miles) 	<ul style="list-style-type: none"> • Both dynamic
Polysemy—degree How rich is Bill Gates?	<ul style="list-style-type: none"> • Total wealth • Ranking among the top rich 	<ul style="list-style-type: none"> • Figure (dollars) • Figure (ranking) 	<ul style="list-style-type: none"> • Both dynamic
Polysemy—degree What is the brightest star visible from Earth?	<ul style="list-style-type: none"> • To naked human eye • To scale 	<ul style="list-style-type: none"> • The sun • Name of instrument 	<ul style="list-style-type: none"> • Both dynamic
Homonymy—acting body Who leads the star ship Enterprise in Star Trek?	<ul style="list-style-type: none"> • Character in the movie • Actor chosen for the role 	<ul style="list-style-type: none"> • Character's name • Actor's name 	<ul style="list-style-type: none"> • Both static
Homonymy—acting body Who is the leader of India?	<ul style="list-style-type: none"> • President of the country • Prime minister in the cabinet system 	<ul style="list-style-type: none"> • Corresponding figure • Corresponding figure 	<ul style="list-style-type: none"> • Both static
Homonymy—acting body How long would it take to get from Earth to Mars?	<ul style="list-style-type: none"> • By space shuttle • By light years 	<ul style="list-style-type: none"> • Corresponding figure • Corresponding figure 	<ul style="list-style-type: none"> • Dynamic • Static
Homonymy—acting body Where does chocolate come from?	<ul style="list-style-type: none"> • Ingredients • Factory 	<ul style="list-style-type: none"> • Related terms (i.e., cacao) • Related terms 	<ul style="list-style-type: none"> • Both dynamic
Homonymy—acting body Who fired Maria Ybarra from her position in San Diego council?	<ul style="list-style-type: none"> • Person • Organization 	<ul style="list-style-type: none"> • Name of person • Name of organization 	<ul style="list-style-type: none"> • Both static

Homonymy—location Where did Bill Gates go to college?	<ul style="list-style-type: none"> • Name of college • Name of city 	<ul style="list-style-type: none"> • Harvard • Boston 	<ul style="list-style-type: none"> • Both static
Homonymy—entity Where is the Taj Mahal?	<ul style="list-style-type: none"> • Monument • Hotel • Restaurant 	<ul style="list-style-type: none"> • Related terms for a relevant domain 	<ul style="list-style-type: none"> • Static • Dynamic • Dynamic

Scope-Based Ambiguity

Origin of Ambiguity Questions	Focus of Plausible Alternatives	(Characteristics of) Prospective Answers
Polysemy—occurrence When did the Jurassic Period end?	<ul style="list-style-type: none"> • Date (period) of the stage • Above PLUS circumstance, characteristics of the occurrence 	<ul style="list-style-type: none"> • Figure; year; related terms such as ago, BC • Above PLUS factors that characterize the end of the period (i.e., disappearance of dinosaurs)
Polysemy—occurrence When did Israel begin turning the Gaza Strip and Jericho over to the PLO?	<ul style="list-style-type: none"> • Date of the occurrence • Above PLUS circumstance of the occurrence 	<ul style="list-style-type: none"> • Date (month/day/year) • Above PLUS relevant descriptions
Polysemy—occurrence When did communist control end in Hungary?	<ul style="list-style-type: none"> • Date of the occurrence • Above PLUS circumstance of the occurrence 	<ul style="list-style-type: none"> • Date (month/day/year) • Above PLUS relevant descriptions
Polysemy—occurrence When did Spain and Korea start ambassadorial relations?	<ul style="list-style-type: none"> • Date of the occurrence • Above PLUS circumstance of the occurrence 	<ul style="list-style-type: none"> • Date (month/day/year) • Above PLUS relevant descriptions
Polysemy—occurrence When did Nixon visit China?	<ul style="list-style-type: none"> • Date (period) of the event • Above PLUS occasion, circumstance of the event 	<ul style="list-style-type: none"> • Date (month/day/year); period • Above PLUS relevant descriptions

II. Syntactic ambiguity

Context-Based Ambiguity

Origin of Ambiguity Questions	Focus of Plausible Alternatives	(Characteristics of) Prospective Answers	Characteristics of Context
Multiple modified terms What costume designer decided that Michel Jackson should only wear one glove?	<ul style="list-style-type: none"> • What-who • What-field 	<ul style="list-style-type: none"> • Name of person • Costume designer's specialty 	<ul style="list-style-type: none"> • Static • Static

Scope-Based Ambiguity

Origin of Ambiguity Questions	Focus of plausible alternatives	(Characteristics of) Prospective Answers
Extent of modified examples What is the name of the rare neurological disease with symptoms such as: involuntary movements (tics), swearing, and incoherent vocalizations (grunts, shouts, etc.)?	<ul style="list-style-type: none"> • Fulfilling all examples (i.e., swearing) listed • Fulfilling part of them 	<ul style="list-style-type: none"> • All examples mentioned • Part of examples mentioned

FEATURE

III. Pragmatic ambiguity

Context-Based Ambiguity

Origin of Ambiguity Questions	Focus of Plausible Alternatives	(Characteristics of) Prospective Answers	Characteristics of Context
Condition of human perception What is the brightest star visible from Earth?	<ul style="list-style-type: none"> • Aided eye • Unaided eye 	<ul style="list-style-type: none"> • Telescope • The sun; a distant star other than the sun with a scale used to measure brightness 	<ul style="list-style-type: none"> • Both dynamic
Time space of inquiry How many Vietnamese were there in the Soviet Union?	<ul style="list-style-type: none"> • In the beginning • In the last stage • During the whole period 	<ul style="list-style-type: none"> • Corresponding figure(s) 	<ul style="list-style-type: none"> • Static • Static • Dynamic
Time space of inquiry How many people does Honda employ in the US? Which country is Australia's largest export market?	<ul style="list-style-type: none"> • As of ongoing year • Based on last year • Based on recent years 	<ul style="list-style-type: none"> • Figure; current year • Figure; last year • Figure; corresponding years 	<ul style="list-style-type: none"> • Dynamic • Static • Static
Domain of regulation What is the legal blood alcohol limit for the state of California?	<ul style="list-style-type: none"> • Operating transportation tools • Engaged in sports activity • Official/nonofficial game 	<ul style="list-style-type: none"> • Related terms (i.e., motor vehicle, aircraft) • Related terms for relevant domain 	<ul style="list-style-type: none"> • All dynamic

Scope-Based Ambiguity

Origin of Ambiguity Questions	Focus of Plausible Alternatives	(Characteristics of) Prospective Answers
Regional space—VS * Where is it planned to berth the merchant ship, Lane Victory, which Merchant Marine veterans are converting into a floating museum? Where is the Keck telescope? Where is Tornado Alley? Where is Microsoft's corporate headquarters located? Where did Dylan Thomas die? Where is the bridge over the river Kwai? Where is Dartmouth College? Where was George Washington born? Where was John Adams born? Where was Lincoln assassinated? Where was Harry Truman born? Where is Ayer's rock? Where is Inoco based? Where did the 6th annual meeting of Indonesia-Malaysia forest experts take place? Where was Ulysses S. Grant born? Where is the actress Marion Davies buried? Where is the Taj Mahal?	<ul style="list-style-type: none"> • Country • State • City • Street • Further specification of location—name of building, street, surrounding entity, etc. 	<ul style="list-style-type: none"> • Country name • State name • City name • Street name • Specific name of place, etc.
Regional space—HS ** What is considered the costliest disaster the insurance industry has ever faced? What is the name of the second space shuttle? Which city has the oldest relationship as a sister-city with Los Angeles? Who was chosen to be the first black chair of the military Joint Chief of Staff?	<ul style="list-style-type: none"> • US only • Any countries including overseas countries 	<ul style="list-style-type: none"> • Related term (i.e., US military) • Terms related to foreign county

<p>Time space—VS When did the Carolingian period begin? When did Muhammad live? When was Dubai's first concrete house built? When was General Manuel Noriega ousted as the leader of Panama and turned over to US authorities? When was the Brandenburg Gate in Berlin built? When did French revolutionaries storm the Bastille? When was the De Beers company founded? When was Microsoft established? When was the San Francisco fire? When was the Triangle Shirtwaist fire? When was China's first nuclear test? When did Nixon die? When did communist control end in Hungary? When was Yemen reunified? When did Jaco Pastorius die? When did Beethoven die? When was the women's suffrage amendment ratified? When did the Vesuvius last erupt? When did Spain and Korea start ambassadorial relations? When did Nixon visit China? When did Lucelly Garcia, a former ambassador of Columbia to Honduras, die? When was London's Docklands Light Railway constructed?</p>	<ul style="list-style-type: none"> • Decade • Year • Month • Day • Time 	<ul style="list-style-type: none"> • Corresponding terms for each
<p>Time space—HS What was the name of the US helicopter pilot shot down over North Korea?</p>	<ul style="list-style-type: none"> • One (particular) helicopter shot down in that region • All helicopters shot down in that region 	<ul style="list-style-type: none"> • No such term/phrase as below • Term/phrase to reflect the extensive scope of the data (i.e., history of, record of)
<p>Time space—HS When was London's Docklands Light Railway constructed?</p>	<ul style="list-style-type: none"> • Range of the construction • Completion of the construction only • Beginning of the construction 	<ul style="list-style-type: none"> • Dates; range of period • Date • Date
<p>Time Space—extent of specified space How much could you rent a Volkswagen bug for in 1966? How much did Mercury spend on advertising in 1993? How much did Manchester United spend on players in 1993? What was the monetary value of the Nobel Peace Prize in 1989?</p>	<ul style="list-style-type: none"> • Amount as of the specified year • Value converted as of the current year 	<ul style="list-style-type: none"> • Monetary amount; measure units • bove PLUS Related descriptions or terms (i.e., conversion, equivalent)
<p>Level of details—reason Why did David Koresh ask the FBI for a word processor?</p>	<ul style="list-style-type: none"> • A simple reason • Detailed explanations 	<ul style="list-style-type: none"> • Reason • Reason and more details
<p>Level of details—terminology What is saltpeter? What is leukemia? What is Head Start? What is a caldera? What is a nematode? What is porphyria? What is a meerkat? What is anorexia nervosa? What are the Valdez principles?</p>	<ul style="list-style-type: none"> • A simple definition • Detailed explanations 	<ul style="list-style-type: none"> • Definition of terminology • Definition and more details
<p>Level of details—product What does the Peugeot company manufacture?</p>	<ul style="list-style-type: none"> • A major area of production • Areas of production • Model name of each product line 	<ul style="list-style-type: none"> • Car • Car, motorcycle, bike • Specific model name

SOURCE—Original sources of questions: TREC 8 and 9 QA Tracks.

NOTE—VS * = Vertical Specificity. HS ** = Horizontal Specificity.