

# HITIQA: A Data Driven Approach to Interactive Question Answering A Preliminary Report

Sharon Small, Nobuyuki Shimizu, Tomek Strzalkowski and Liu Ting

ILS Institute  
The State University of New York at Albany  
1400 Washington Avenue  
Albany, NY 12222  
{small,ns3203,tomek,tl7612}@albany.edu

## Abstract

HITIQA is an interactive question answering technology designed to allow intelligence analysts and other users of information systems to pose questions in natural language and obtain relevant, factual answers, or the assistance they require in order to perform their tasks. Furthermore, our objective in HITIQA is to allow the user to submit exploratory, analytical, non-factual questions, such as “*What has been Russia’s reaction to U.S. bombing of Kosovo?*” HITIQA uses interactive natural language dialogue to score passages of the retrieved documents in relation to the query. The information obtained from this dialogue will aid the system in clarifying the user’s query, retrieving additional relevant documents and generating succinct answers to analytical questions. This paper will present a preliminary series of results and the results of a pilot evaluation by NIST.

## Introduction

Analytical questions such as, “*What recent disasters occurred in tunnels used for transportation?*” are a unique and challenging problem beyond the ability of traditional question answering systems. This example question would possibly require a list or a table of appropriate facts organized according to an analyst’s instructions. Clearly, this cannot be an isolated exchange; in most cases the system must engage the analyst in a dialogue to clarify the intentions and goals of the question.

HITIQA project is part of the ARDA AQUAINT program that aims to make significant advances in the state of the art of automated question answering. In this paper we focus on two aspects of our work:

1. Question Semantics: how the system “understands” user requests.
2. Human-Computer Dialogue: how the user and the system negotiate this understanding.

We will also present our evaluation results from a series of pilot tests run by NIST. This is the first known testing of a dialogue question answering system. This testing was aimed at evaluating the dialogue component of the system.

## Document Retrieval

When the user poses a question to a system sitting atop a huge database of unstructured data (text files), the first order of business is to reduce that pile to perhaps a handful of documents where the answer is likely to be found. This means, most often, document retrieval, using fast but non-exact selection methods.

The objective in HITIQA is to allow the user to submit exploratory, analytical, non-factual questions. Because of the variety of syntactic forms that can be used to form such questions, with only a loose connection between their syntax and the user intent, the questions are only minimally processed prior to initial document retrieval. Their semantics is more likely to be deducted from the documents returned than through a detailed analysis of their particular form. The detailed question processing that has been successful in understanding of factual questions (Harabagiu et al, 2000; Hovy et al, 2000) is clearly less beneficial here. Instead, questions are simply tokenized and sent to SMART for document retrieval. In some cases it may be beneficial to initiate a dialogue with the user to clarify queries that the system has judged to be ambiguous. For example:

User: "Please give me biographical information on Elizardo Sanchez."

HITIQA: "What do you mean by biographical information?"

User: "his childhood, education and early involvement on opposing the Cuban government"

**Example 1:** Interactive Query Clarification Dialogue between user and HITIQA

This information can now be added to the query and used in answer generation. HITIQA currently focuses on dialogue after document retrieval. The Interactive Query Clarification is in development.

We are retrieving the top fifty documents from three gigabytes of newswire (AQUAINT corpus plus web-harvested documents). The system segments these documents into paragraphs and filters out duplicates. The user's question is then translated into a GoalFrame for later use in the assessment of retrieved data and by the dialogue manager. The concept of a GoalFrame is discussed in detail later.

## Clustering

The retrieved data is our best source of information to understand the "meaning" of the user's query. Therefore, the system must "understand" this data to the best of its ability. N-gram clustering and concept generation are used to determine the main topics returned for the user's query.

Retrieved documents are first broken into naturally occurring paragraphs. Duplicate paragraphs are filtered and these resulting passages are clustered. Typically three to six clusters are generated (for details of the clustering algorithm see Hardy et al, 2002a).

HITIQA uses concepts to help "understand" the topics of the retrieved data. Concepts are generated for each cluster using that cluster's seed paragraphs. These seed paragraphs are first tagged using a part of speech tagger. HITIQA uses the noun phrases from the tagging process and JWNL<sup>1</sup> (Java Wordnet<sup>2</sup>

Library) to generate the base concepts. This process looks for a common relation between the noun phrases' hypernym trees. Using the concept generator, a paragraph with the nouns "rifle" and "machine gun" would return "weaponry" as one of its concepts, as shown in Figure 1.

```
rifle
=> firearm, piece, small-arm
=> gun
=> weapon, arm, weapon system
=> weaponry, arms...

machine gun
=> firearm, piece, small-arm
=> gun
=> weapon, arm, weapon system
=> weaponry, arms...
```

**Figure 1:** A portion of Wordnet hypernym trees for "rifle" and "machine gun"

These concepts, generated from the clusters, give the system its first glimpse into the possible meanings of the user's query. The concepts are used during the Framing process and again during the Clarification Dialogue as described below.

## Framing

HITIQA uses *text framing* to narrow the gap between the user's question and the system's "understanding" of this question, namely the retrieved data. The framing is an attempt to impose a partial structure on the text that would allow the system to more easily compare different text pieces and also to communicate with the user about the retrieved data.

The GoalFrame is the first frame built. It is generated from the user's question using a modified version of GATE<sup>3</sup>. The most recent version of GATE tags locations, organizations, people etc. We have modified GATE to separate organizations into companies and organizations, and we have also expanded GATE by tagging new concepts such as industries. The GoalFrame generated from the question, "What information is available on Iraq's amassing of weapons of mass destruction and the United Nations' inspections?" is shown in Figure 2 below.

<sup>1</sup> JWNL is an API for accessing Wordnet-style relational dictionaries developed at SourceForge.net.

<sup>2</sup> Wordnet is a lexical database developed at Princeton University (Miller, 1995).

<sup>3</sup> GATE is Generalized Architecture for Text Engineering, an information extraction system developed at the University of Sheffield (Cunningham, 2000).

Targets: [mass destruction, inspections, weapons, amassing]

...

Locations: [Iraq]

Organizations: [United Nations]

**Figure 2:** HITIQA generated GoalFrame

The GoalFrame consists of a list of attribute-value pairs. All the factual information in the GoalFrame comes from the question. Attribute values, except for the Target, are obtained using specialized text extractors in our modified version of GATE. Targets are filled with qualifying phrases from the query.

Data frames are built from the retrieved data, after clustering into several topical groups. Since clusters are built out of small text passages, we associate a frame with each passage that serves as a seed of a cluster. We merge paragraphs prior to framing based on simple anaphor resolution. The main advantage of working with small text pieces is that *framing* them (i.e., mapping them onto a frame) becomes a relatively less complicated process – in particular there are fewer ambiguities and uncertainties normally associated with longer texts. In determining the target of these data frames, we match these text segments against the targets of the GoalFrame. If there is an exact match against a GoalFrame target, this becomes the data frame's target. If more than one match is found they become the sub-targets of the data frame. It is also possible for there to be no match against the GoalFrame target. In this situation we choose the target from the list of generated concepts for the corresponding cluster. An example data frame generated from the text retrieved in response to the query about Iraq is shown in Figure 3 below. After the initial framing is done, frames judged to be related to the same concept or event are merged and values of their attributes are combined.

target: mass destruction

sub-target: [weapons]

location: [Kuwait, Iraq, Iraq, Baghdad]

companies: []

people: []

organization: [U.N., Security Council, U.N.]

document date: [1998, 6, 13]

source: Baghdad, Iraq(AP)-

text: [The chief U.N. weapons inspector began talks Saturday with top Iraqi officials on steps that **Baghdad** must take to convince arms monitors that **Iraq** has destroyed its

weapons of mass destruction.

The U.N. inspectors certification is required for the **Security Council** to consider lifting sweeping economic sanctions imposed after **Iraq's** 1990 invasion of **Kuwait**.]

Relevance: Matches on all elements found in goal frame: [location, organization, target=[mass destruction, inspections, weapons, amassing]]

**Figure 3:** An HITIQA generated data frame. Words in bold were used to fill the Frame.

### Judging Frame Relevance

We judge a particular data frame as relevant, and subsequently the corresponding segment of text as relevant, by comparison to the GoalFrame. The data frames are scored based on the number of conflicts found between itself and the GoalFrame. If a frame is found to have no conflicts, it is scored as relevant for the given query, and given a conflict score of zero. All other data frames are scored with an incrementing conflict value, one for frames with one conflict with the GoalFrame, two for two conflicts etc. Frames that conflict with all information found in the query are given a score of 99 indicating that it is irrelevant and not to be used in any further processing. The frame in Figure 3 is scored as relevant to the query (0 conflicts).

### Enabling Dialogue with the User

Framed information allows HITIQA to conduct a meaningful dialogue with the user. The purpose of the dialogue is to help the user to navigate the answer space and to solicit from the user more details as to what information he or she is seeking. The main principle here is that the dialogue is at the information semantic level, not at the information organization level. Thus, it is okay to ask the user whether information about AIDS conference in Cape Town should be included in the answer to a question about combating AIDS in Africa. However, the user should never be asked if a particular keyword is useful or not, or if a document is relevant or not. We have developed a 3-pronged strategy:

1. Narrowing dialogue: ask questions that would allow the system to reduce the size of the answer set.
2. Expanding dialogue: ask questions that would allow the system to decide if the an-

swer set needs to be expanded by information just outside of it (near-misses).

3. Fact seeking dialogue: allow the user to ask questions seeking additional facts and specific examples, or similar situations.

Of the above, we have thus far implemented the first two options as part of the preliminary clarification dialogue. The clarification dialogue is when the user and the system negotiate the task that needs to be performed. We can call this a “triaging stage”, as opposed to the actual problem solving stage (point 3 above). In practice, these two stages are not necessarily separated and may be overlapping throughout the entire interaction. Nonetheless, these two have decidedly distinct character and require different dialogue strategies on the part of the system.

Our approach to dialogue in HITIQA is modeled to some degree upon the mixed-initiative dialogue management we adopted in another project (AMITIES; Hardy et al, 2002b). The main advantage of the AMITIES model is its reliance on data-driven semantics which allows for spontaneous and mixed initiative dialogue to occur.

By contrast, the major approaches to implementation of dialogue systems to date rely on systems of functional transitions that make the resulting system much less flexible. In the grammar-based approach, which is prevalent in commercial systems, such as Nuance’s various telephony products (<http://www.nuance.com>), as well as in practically oriented research prototypes; e.g., systems developed under the DARPA Communicator program (<http://www.darpa.mil/iao/Communicator.htm>), or Rochester TRAINS system (Ferguson & Allen, 1998) a complete dialogue transition graph is designed to guide the conversation and predict user responses, which is suitable for closed domains only. In the statistical variation of this approach, a transition graph is derived from a large body of annotated conversations. This latter approach is facilitated through a dialogue annotation process, e.g., using Dialogue Act Markup in Several Layers (DAMSL) (Allen and Core, 1997), which is a system of functional dialogue acts.

Nonetheless, an efficient, spontaneous dialogue cannot be designed on purely functional layer, therefore we are here primarily interested in the semantic layer, that is, the information exchange and information building effects of a conversation. In order to properly understand a dialogue, both semantic and functional layers need to be considered. In this paper

we are concentrating exclusively on the semantic layer.

### Clarification Dialogue

Data frames with a conflict score of zero form the initial answer space. Depending upon the size of this set and the presence of other frames outside of it, the system either proceeds to generate the answer or initiates a dialogue with the user. For example, if the answer space appears too large, the system may ask the user how to narrow it. The presence of large groups of near-miss frames (with 1 or 2 conflicts with the GoalFrame) may indicate potentially relevant information, and the user will be consulted for possible broadening of the question. HITIQA uses Clarification Dialogues (CD) to determine the relevance of the non-zero frames. Currently, we only generate Clarification Dialogue for one-conflict frames.

A one-conflict frame has only a single attribute mismatch with the GoalFrame. If the GoalFrame does not have a target (i.e., question is too vague), all data frames are scored as conflicting with the target of the GoalFrame. When the GoalFrame target is empty, HITIQA will only begin a CD on one-conflict frames that have a target that is related to the GoalFrame target or on one-conflict frames when the GoalFrame target is empty.

User: “*Who is Elizardo Sanchez?*”

HITIQA: “*Are you interested in seeing information about civil rights as it is related to Elizardo Sanchez?*”

One-Conflict Frame:

target: *civil rights*

sub-target: []

location: [*Cuba, Cuba, Cuba, Cuba*]

companies: []

people: [*Gonzalez, Alejandro, Sanchez, Sanchez*]

value: []

percent: []

organization: [*Foreign Ministry*]

date: [*past decade, past year, Tuesday*]

document date: [*2000, 1, 11*]

source: *HAVANA (AP) –*

text: [**Foreign Ministry** spokesman **Alejandro Gonzalez** did not immediately return a call seeking comment **Tuesday**. The **Cuban** government's position is that it holds no political prisoners, only common criminals.

*“I consider that the situation for civil and political rights in **Cuba** has worsened over the **past year**... owing to that **Cuba** continues to be the only closed society in this hemi-*

sphere," **Sanchez** said. "There have been no significant release of prisoners, the number of people sanctioned or processed for political motives increased. During the year, there took place the widest waves of repression over the **past decade**," he concluded. **Sanchez**, who himself spent many years in **Cuban** prisons, is among the communist island's best known opposition activists. The commission he heads issues a report on civil rights every six months, along with a list of people it considers to be imprisoned for political motives. ]

**Example 2:** An actual Clarification Dialogue and one of the Frames that was used in generating this dialogue. Words in bold were used to fill the Frame.

During the dialogue, as new information is obtained from the user, the GoalFrame is updated and the scores of all the data frames are reevaluated. The system may interpret the new information as a positive or negative. Positives are added to the GoalFrame. Negatives are stored in a NegativeGoalFrame and will also be used in the re-scoring of the data frames, possibly causing conflict scores to increase. The NegativeGoalFrame is created when HITIQA receives its first negative response from the user. The NegativeGoalFrame includes information that HITIQA has identified as being of no interest to the user. If the user responds the equivalent of "yes" to the dialogue in Example 2, *civil\_rights* will be added to the target list in GoalFrame and all one-conflict frames with a *civil\_rights* target will be rescored to Zero conflicts, two-conflict frames with *civil\_rights* as a target will be rescored to one, etc. If the user responds "no", the NegativeGoalFrame will be generated and all frames with *civil\_rights* as a target will be downgraded to 99 frames.

The clarification dialogue will continue until all the significant sets of one-conflict frames are either included in the answer space (through user broadening the scope of the question that removes the initial conflicts) or dismissed as not relevant. When HITIQA reaches this point it will re-evaluate the data frames in its answer space. If there are too many frames (more than a pre-determined upper threshold), the dialogue manager will offer to the user to narrow the question, thus reducing the number of frames judged relevant. On the other hand, if the size of the answer space is still too small (i.e., too few frames), the dialogue manager will suggest to the user ways of further broadening the question, thus making more data frames relevant, or possibly retrieving new documents by adding terms learned

in the Clarification Dialogue to the query. When the number of frames is within the acceptable range, HITIQA will generate the answer using the text from the frames in the current answer space. The user may end the dialogue at any point and have an answer generated given the current Frames.

### Narrowing Dialogue

HITIQA attempts to reduce the number of frames judged to be relevant through a Narrowing Dialogue. This is done when the answer space contains too many elements to form a succinct answer. This typically happens when the initial question turns out to be too vague or unspecific, with respect to the available data. For example, the Iraq question could be narrowed using the date attribute as a guide:

*"We have found a large amount of information judged to be relevant and would like to narrow your results. Are you interested in seeing this information only for specific years? (Valid years for the current query are 1998, 1999, 2001, and 2002.)"*

**Example 3:** Example of a HITIQA generated Narrowing Dialogue question

The user may wish to see only the most recent information and ask for 2002, in which case the system would base its answer on the frames with date attribute set to 2002.

### Broadening Dialogue

The system can also attempt to increase the number of frames judged relevant through a Broadening Dialogue (BD) whenever the answer space appears too narrow, i.e., contains too few zero-conflict frames. The exact definition of when the answer space is too small depends on multiple factors, including the question itself. We are conducting experiments to learn more about this. Furthermore, the BD will only occur if there are one-conflict frames. For example, in the case of the Iraq question, several frames were judged to have a conflict with the GoalFrame on the organization attribute (here, *the United Nations*), but matched on all other attributes. In these near-miss frames, the organization was the *UNSCOM (United Nations Special Commission)*, which the system lacked knowledge to associate with the U.N. Therefore, the system asked the user whether they were also interested in information involving *UNSCOM* and possibly other organizations. Broadening ques-

tions can be asked on any of the attributes which have values in the GoalFrame.

### Interpreting User Responses

While the system poses yes-no questions to the user in the dialogue, in addition to a simple yes or no response, the user often supplies additional information, in order to help guide the system, as can be seen in the following examples.

*“No, not necessary. I need his background information, such as his early childhood, his education, any early involvement in opposing the Cuban government.”*

*“Yes, if it is related to his background.”*

*“Yes, also per capita income and stock market value will help.”*

*“Can you give information on when they were formed?”*

**Example 4:** Actual user responses during AQUAINT evaluations

In order to exploit this additional information, the user’s response is scored against their initial query. The following pseudo code shows our strategy of analyzing user responses.

```
if(user response contains yes or no)
    if (the response is not similar to query,
        and there are added terms)
        then edit the Goal Frame with
            new terms.
    go on with the dialogue
else
    if(the response is a question itself
        “what is/are X?”)
        then show the definition of X
            using WordNet.
    else
        edit the Goal Frame with the new terms.
```

**Figure 4:** Response analyzer pseudo code

The generic terms/phrases common in the response, such as “general information” and “help”, are considered “response stop words” and not added to the Goal Frame. Words identified as new information are added to the goal frame as seen in Example 5.

User: *“Yes, also per capita income and stock market value will help.”*

HITIQA: *“Thank you, capita income, and stock market value have been added to your goal frame.”*

**Example 5:** HITIQA action for user’s response with extra information

### No Answer/Answer

HITIQA also has the ability to realize when there is no data retrieved that will satisfy the user’s query. It is able to generate an informational message letting the user know exactly where the data does not meet the query needs. This situation is recognized when all frames generated have a score of two conflicts and higher.

User: *“What information is available on the cellular telephone network in the Ivory Coast?”*

...

HITIQA: *“We have no data on the combination of cellular telephone network and the Ivory Coast. Specifically there is nothing on the cellular telephone network.”*

**Example 6:** A “No Answer” Dialogue

Currently, HITIQA has done no processing on the text after the dialogue has finished. The text of frames with a score of zero conflicts are ordered by date and outputted to the user. Typically the answer to these analytical type questions will require many pages of information. Example 7 below shows the first portion of the answer generated by HITIQA for the Iraq sample query using just the newswire data for the years 1998 through 2000. Current work is focusing on improving the generated answer even further.

### 1999:

*Saeed H. Hasan, Iraq's Ambassador to the United Nations made the statement during an interview with Xinhua.*

*Hasan said the UNSCOM and its Executive Chairman Richard Butler must go. The Ambassador said UNSCOM, in charge of dismantling Iraq's weapons of mass destruction, conducted their work in a unjust, unfair way which prolonged the U.N. sanctions against Iraq. Hasan said Iraq has complied with relevant U.N. resolutions and destroyed all weapons of mass destruction, but UNSCOM failed to report this fact.*

*The Ambassador said UNSCOM should show the Security Council evidence when It doubted that Iraq was hiding weapons, but It failed to offer any.*

*Hasan said signs have shown that UNSCOM, which should work for the United Nations, has become a tool of*

a specific country and only serves the interests of this country.

He urged the United Nations Security Council to have it abolished.

**1998:**

*Iraq and the United Nations are locked in a standoff over weapons inspections which the country is compelled to unconditionally accept under the terms that ended the 1991 Gulf War over Kuwait.*

*Iraq said Wednesday the lifting of sanctions is the only way to defuse its current dispute with the United Nations over weapons inspections.*

*Earlier on Thursday the White House said Iraq's decision to freeze weapons inspections was unacceptable and vowed to keep pressure on Iraq unless it complies with the U.N. resolution. "This noncompliance is completely unacceptable," White House spokesman P.J. Crowley said.*

**Example 7:** Answer generated by HITIQA to the Iraq query.

**Evaluations**

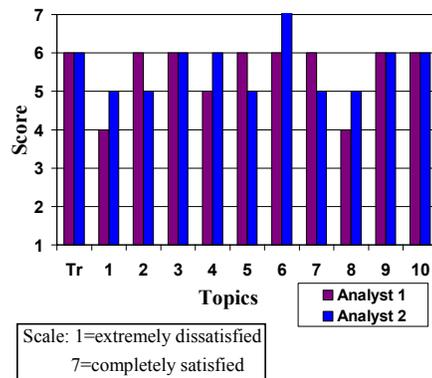
We have just completed the first round of a pilot evaluation for testing the interactive dialogue component of HITIQA. The purpose of this first stage of evaluation is to determine what kind of dialogue is acceptable/tolerable to the user and whether an efficient navigation through the answer space is possible. HITIQA was blindly tested by two different analysts on eleven different topics. The analysts were given complete freedom in forming their queries and responses to HITIQA's questions. They were only provided with descriptions of the eleven topics the systems would be tested on. The analysts were given 15 minutes for each topic to arrive at what they believed to be an acceptable answer. During testing a Wizard was allowed to intervene if HITIQA generated a dialogue question/response that was felt inappropriate. The Wizard was able to override the system and send a Wizard generated question/response to the analyst. HITIQA Wizard intervened an average of 13% of the time.

**Results of Evaluation**

This results are for information purposes only as it was not a formal evaluation. HITIQA earned an average score of 5.8 from both Analysts for dialogue. The highest score possible was a 7 for each dialogue. The Analysts were asked to grade each scenario for success or failure. We divide the failures into three categories:

- 1) the user gives up on the system for the given scenario
- 2) the 15 minute time limit was up
- 3) the data was simply not in the database

**AQUAINT DIALOGUE EVALUATIONS**



**Figure 5:** Graph of HITIQA Dialogue Scores from the AQUAINT Evaluation

**Topics**

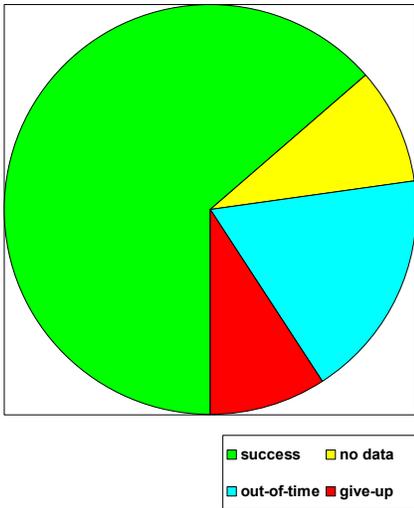
|                            |
|----------------------------|
| Tr = Sherron Watkins       |
| 1 = Japanese Joint Venture |
| 2 = Indonesia economics    |
| 3 = Black Sea pollution    |
| 4 = AIDS in Africa         |
| 5 = Afghanistan opium      |
| 6 = Elizardo Sanchez       |
| 7 = Microsoft viruses      |
| 8 = Ivory Coast cellular   |
| 9 = FARC                   |
| 10 = Robotic Surgery       |

**Figure 6:** AQUAINT dialogue evaluation topics

HITIQA had a 63% success rate for Analyst 1 and a 73% success rate for Analyst 2. Topic 6 was purposely left without data in our database, in order to evaluate how the analyst would respond to the system being unable to find any relevant data. It is interesting to note that when the system told the ana-

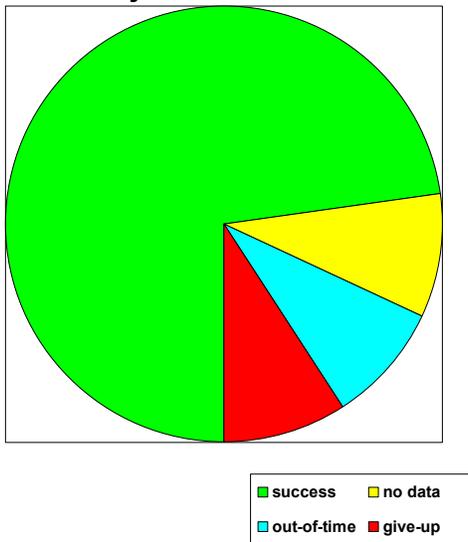
lyst that there was no data available for a given topic, both analysts then graded this as a failure, even though it was the proper response from the system. We believe this may have been caused by the analysts being able to distinguish between the system no being able to find the information and the information not being available to the system.

**Analyst 1 Success Rate**



**Figure 7: Results of Analyst 1**

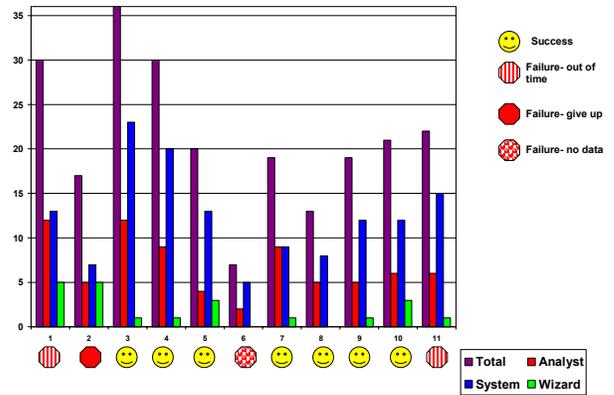
**Analyst 2 Success Rate**



**Figure 8: Results of Analyst 2**

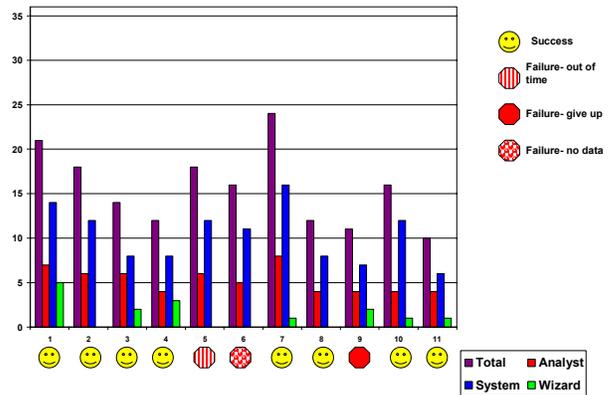
Figure 9 and Figure 10 show the breakdown of total dialogue utterances. This is broken down between analyst utterances, system generated utterances and wizard utterances. It is interesting to note that, the percentage of total utterances for both analysts was roughly the same. The percentage for Analyst One was 32% of the total, while the percentage for Analyst Two was 34% of the total.

**Breakdown of dialogue utterances Analyst One**



**Figure 9: Breakdown of utterances: Analyst One**

**Breakdown of dialogue utterances Analyst Two**



**Figure 10: Breakdown of utterances: Analyst Two**

**Future Work**

This paper describes a work in progress. We expect that the initial specification of content frame will evolve as we subject the initial system to more demanding evaluations. Currently, the frames are not topically specialized, and this appears the most logical next refinement, i.e., develop several (10-30)

types of frames covering different classes of events, from politics to medicine to science to international economics, etc. This is expected to increase the accuracy of the dialogue as is the interactive visualization which is also under development. Answer generation will involve fusion of information on the frame level, and is currently in an initial phase of implementation.

### Acknowledgements

This paper is based on work supported by the Advanced Research and Development Activity (ARDA)'s Advanced Question Answering for Intelligence (AQUAINT) Program under contract number 2002-H790400-000.

### Appendix

This appendix contains two complete sample dialogues from the NIST pilot evaluation.

**Dialogue One:** This is the third dialogue interaction in a series of eleven dialogues for Analyst One.

**Analyst:** "I am researching business joint ventures with the Japanese. Can you name for me any new joint business venture involving a Japanese company. I need the name of the participants and the name and activity of that new company."

**HITIQA:** "We are retrieving documents now. Please wait..."

**HITIQA:** "We are processing retrieved documents. Please wait..."

**HITIQA:** "We are beginning your answer search. Please wait..."

**HITIQA:** "We have found relevant information for your query. (You may ask for answer generation at any point now.)

*We have also found other information that may be relevant. Are you interested in seeing information about transactions as it is related to Japanese joint ventures?"*

**Analyst:** "Yes."

**HITIQA:** "Thank You. We have added transaction as relevant.

*We have also found other information that may be relevant. Are you interested in seeing general information on company as it is related to Japanese joint ventures?"*

**Analyst:** "I specifically need to know the names of the participants, the name of the new business joint ventures and the activity of that company."

**HITIQA:** "We are generating your answer."

**HITIQA:**

1999:

*BRUSSELS, Belgium (AP) - The European Commission gave the go-ahead Thursday for leading telecommunications companies AT&T and British Telecommunications to take a combined 30 percent stake in Japan Telecom Co.*

*The Commission, the European Union's executive arm, said that "given the relatively modest presence of Japan Telecom in the Japanese domestic telecommunications sector, and in the international telecommunications services provided from Japan, the transaction does not raise any particular competition concerns." Japan Telecom is Japan's third-largest long-distance and international carrier. It also has stakes in digital mobile companies.*

*AT&T is the largest long-distance and international carrier in the United States, while BT is Britain's biggest telecom company. Under the deal, AT&T and BT will each hold a 15 percent stake in Japan Telecom.*

*The three firms are planning a broad alliance in which AT & T and BT will integrate most of their Japanese operations with Japan Telecom's units. In addition, the two companies will dispatch executives to the Japanese firm.*

**HITIQA:** "We have additional information like this, would you like to see it?"

**Analyst:** "Can you tell me which of these was the first joint venture?"

**HITIQA:** "We are processing your query."

**Analyst:** "forget that last question"

**Analyst:** "Can you give me the additional info on joint ventures that you referenced."

**HITIQA:**

1999:

*Softbank will have a 52.5 percent stake in the venture, which will have capital of 200 million yen (\$1.6 million). The second biggest partner, Japanese toy wholesaler Happinet Corp., will hold a 17.5 percent stake. Bandai Corp., the maker of Tamagotchi toys, also will be involved.*

*Using the convenience store is designed to overcome problems that have kept computer commerce from becoming popular in Japan: the low level of credit card use and concerns about security. Separately, Softbank President Masayoshi Son confirmed that his company will establish a 230 billion yen (\$1.89 billion) venture capital fund to invest in Internet start-ups.*

*More than 95 percent of the fund will be invested in U.S. companies, with the balance in Japanese firms. Softbank itself will supply 90 billion yen (\$738 million) of the venture capital. Softbank, a software wholesaler and computer-magazine publisher, has become increasingly involved in Internet businesses.*

*AT&T is acquiring the IBM Global Network which specializes in data networking and Internet services for large companies."*

**Analyst:** "Success"

**Dialogue Two:** This is the third dialogue interaction in a series of eleven dialogues for Analyst Two.

**Analyst:** "Want current and potential treatments for attacking the AIDS epidemic in Africa."

**HITIQA:** "We are retrieving documents now. Please wait..."

**HITIQA:** "We are processing retrieved documents. Please wait..."

**HITIQA:** "We are beginning your answer search. Please wait..."

**HITIQA:** "We have found relevant information for your query. (You may ask for answer generation at any point now.)

*We have also found other information that may be relevant. Are you interested in seeing information about The International AIDS Conference as it is related to AIDS in Africa?"*

**Analyst:** "Only to answer the questions on treatment for AIDS. Can you give me the types of treatments currently used in Africa?"

**HITIQA:** "Thank you. Are you interested in seeing information about The World Health Organization as it is related to AIDS in Africa?"

**Analyst:** "If it gives me some treatments, then that will be good."

**HITIQA:**  
2000:

*Sandra Thurman, head of the White House AIDS Office, also urged other countries and international organizations to step up their efforts as well.*

*"The United States cannot do it alone," she said at a news conference at the 13th International AIDS Conference being held in Durban, South Africa.*

*Chirac's speech was delivered by the country's Health Minister Dominique Gillot at the 13th International AIDS Conference which opened in Durban, South Africa, on Sunday.*

*Fallout from the AIDS crisis is changing population structures in ways that will rip the social and economic fabric, according to the report released by the Joint U.N. Program on HIV/AIDS (UNAIDS). A total of 13.2 million children have lost their mother or both parents to the disease since the epidemic began.*

**HITIQA:** "We have additional information like this, would you like to see it?"

**Analyst:** "Please try to search on possible treatments such as AZT."

**HITIQA:** "We are retrieving documents now. Please wait..."

**HITIQA:** "We are processing retrieved documents. Please wait..."

**HITIQA:** "We are beginning your answer search. Please wait..."

**HITIQA:**

*2000: While the panel meets, South Africa continues its campaign to stop the spread of the disease, find humane treatments for those infected and help in the push for an AIDS vaccine.*

*"There is no substance to the allegation that there is any hesitation on the part of our government to confront the challenge of HIV/AIDS," he said.*

*The South African government has also come under criticism for refusing to finance a program to treat pregnant mothers with the anti-AIDS drug AZT, and Mbeki has said there was significant evidence the treatment was so dangerous it would do more harm than good.*  
*1999: An estimated 30 percent of women in the urban areas of sub-Saharan countries, such as Uganda, are infected with HIV, said Fleming. The majority of those women receive no treatment for HIV, and about one-third of infected women pass the virus to their babies, either in utero or through breast-feeding."*

**HITIQA:** "We have additional information like this, would you like to see it?"

**Analyst:** "Yes."

**HITIQA:** "1999: "This study represents the most promising advance to date toward the goal of finding strategies that can be used worldwide to prevent the spread of HIV from HIV-infected mothers to their infants," said Dr. Anthony Fauci, director of the National Institute of Allergy and Infectious Diseases, which funded the study.

*Equally important, researchers said, is the new regimen's simplicity. In contrast to existing drug treatments that must be given twice daily for weeks or months, the new therapy involves giving a single oral dose of an antiviral drug called nevirapine to women in labor \_ the first time in pregnancy that many women in poor countries see a health professional.*

*Within three days of birth, her infant gets a single oral dose of the drug. An added bonus is that the drug doesn't require refrigeration.*

*"We probably have 800,000 or even a million infected kids being born ever year," said Dr. Brooks Jackson of Johns Hopkins University, who led the Uganda study. "If this regimen could be used worldwide, that's up to 500,000 lives saved a year." Crispus Kiyonga, Uganda's health minister, said the results "provide real hope that we may be able to protect many of Africa's next generation from the ravages of AIDS." Dr. John Sullivan of the University of Massachusetts Medical School, who first suggested nevirapine's use to prevent mother-infant transmission in 1992, predicted the regimen will be "one of the most important breakthroughs since the AIDS virus was isolated.*

**Analyst:** "Success"

## References

- (Allen and Core, 1997). J. Allen and M. Core. 1997. Draft of DAMSL: Dialog Act Markup in Several Layers.  
<http://www.cs.rochester.edu/research/cisd/resource/damsl/>.
- (Bagga et al, 2000). Bagga, A., T. Strzalkowski, and G.B. Wise. (2000). PartsID: A Dialog-Based System for Identifying Parts for Medical Systems. Proc. of the ANLP-NAACL-2000.
- (Cunningham et al, 2000). Cunningham, H., D. Maynard, K. Bontcheva, V. Tablan and Y. Wilks. (2000) Experience of using GATE for NLP R&D. In Workshop on Using Toolsets and Architectures To Build NLP Systems at COLING-2000, Luxembourg.
- (Ferguson & Allen, 1998). George Ferguson and James Allen, "TRIPS: An Intelligent Integrated Problem-Solving Assistant," in *Proceedings of the Fifteenth National Conference on Artificial Intelligence (AAAI-98)*, Madison, WI, 26-30 July 1998, pp. 567-573.
- (Hardy et al, 2002a). H. Hardy, N. Shimizu, T. Strzalkowski, L. Ting, B. Wise and X. Zhang (2002). Cross-Document Summarization by Concept Classification. Proceedings of SIGIR-2002, Tampere, Finland.
- (Hardy et al, 2002b). H. Hardy, K. Baker, L. Devillers, L. Lamel, S. Rosset, T. Strzalkowski, C. Ursu and N. Webb. 2002. Multi-layer Dialogue Annotation for Automated Multilingual Customer Service. ISLE Workshop, Edinburgh, Scotland.
- (Harabagiu et al, 2000). Harabagiu, S., M. Pasca and S. Maiorano. (2000) Experiments with Open-Domain Textual Question Answering. In Proceedings of COLING-2000, pages 292--298.
- (Hovy et al, 2000). Hovy, E., L. Gerber, U. Hermjakob, M. Junk, C-Y. Lin (2000). Question Answering in Webclopedia. Notebook Proceedings of Text Retrieval Conference (TREC-9).
- (Miller, 1995). Miller, G.A. 1995. WordNet: A Lexical Database. *Communication of the ACM*, 38(11):39--41.