AQUA: Asking Questions and Understanding Answers

Ashwin Ram

Yale University
Department of Computer Science
New Haven, CT 06520-2158

Abstract

Story understanding programs are often designed to answer questions to demonstrate that they have adequately understood a story (e.g., [Leh78]). In contrast, we claim that *asking* questions is central to understanding. Reading a story involves the generation of questions, which in turn focus the understander on the relevant aspects of the story as it reads further.

We are interested in the kinds of questions that people ask as they read. In this paper, we talk about the origin of these questions in the basic cycle of understanding, and their effect on processing. We present an understanding algorithm based on our theory of questions, which we have implemented in a computer program called AQUA (Asking Questions and Understanding Answers).

I. Question-driven understanding

"The students seemed to understand the lecture — at least they were asking the right questions." — A teacher.

When we read a story, we are constantly trying to relate the events in the story to what we already know. We build motivational and causal explanations for the events in the story in order to understand why the characters acted as they did, or why certain events occurred or did not occur. The central claim of this paper is that an understander asks *questions* in order to understand the story, build explanations for it, and integrate it into memory. The depth of understanding that the understander achieves depends on the questions that it asks.

Consider, for example, the following excerpt from a rather unusual story which appeared on the front page of the New York Times a couple of years ago:

> **Boy Says Lebanese Recruited Him as Car Bomber.**
>
> New York Times, Sunday, April 14, 1985.
>
> JERUSALEM, April 13 — A 16-year-old Lebanese was captured by Israeli troops hours before he was supposed to get into an explosive-laden car and go on a suicide bombing mission

to blow up the Israeli Army headquarters in Lebanon. ...

> What seems most striking about [Mohammed] Burro's account is that although he is a Shiite Moslem, he comes from a secular family background. He spent his free time not in prayer, he said, but riding his motorcycle and playing pinball. According to his account, he was not a fanatic who wanted to kill himself in the cause of Islam or anti-Zionism, but was recruited for the suicide mission through another means: blackmail.

The premise is that reading involves the generation and transformation of questions. This story was read out loud to a class of graduate students. As they heard the story, the students voiced the questions that occurred to them. Here are a few of the questions that they asked:

1. Why would someone commit suicide if he was not depressed?
2. Did the kid think he was going to die?
3. Are car bombers motivated like the Kamikaze?
4. Does pinball lead to terrorism?
5. Who blackmailed him?
6. What fate worse than death did they threaten him with?
7. Why are kids chosen for these missions?
8. Why do we hear about Lebanese car bombers and not about Israeli car bombers?
9. Why are they all named Mohammed?
10. How did the Israeli know where to make the raids?
11. How do Lebanese teenagers compare with U.S. teenagers?

Some of the questions seem pretty reasonable, (e.g., *Did the kid think he was going to die?*), but some are rather silly in retrospect (e.g., *Does pinball lead to terrorism?*). Some, though perfectly reasonable questions, aren't central to the story itself, but instead relate to other things that the person concerned was reminded of, things that he was wondering about or interested in (e.g., *Why do we hear about Lebanese car bombers and not about Israeli car bombers?*).

A. The nature of questions

Questions such as the above arise naturally during the course of understanding. Let us summarize our central claims about the nature of such questions.

- Since the ultimate goal of understanding is the integration of new input with what the system already knows, questions that arise during the integration process represent *difficulties* in processing.

- The understander needs to ask these questions in order to perform explanation tasks effectively. *Asking the right questions is central to achieving a greater depth of understanding*. For example, thinking about the Kamikaze question (3) is likely to lead to a better understanding of the boy's motivations than is thinking about the pinball question (4).

- Since questions arise from unusual input for which the understander does not have the appropriate processing structures in memory, or from explicit contradictions between the predictions supplied by memory structures and the actual input, *questions reflect that part of the input that needs extra attention, i.e., that part of the input that the understander ought to focus on*. In other words, questions represent what the understander is interested in finding out with respect to its goal of understanding the story. They should be used to drive the understanding process.

- *The process is dynamic* in that new input generates new questions or transforms old ones, which in turn affects further processing of the story and of future stories.

B. Research issues

We are approaching the problem of story understanding as a process involving the generation and transformation of questions. We are designing a computer program that asks creative questions while it reads a story in order to raise the level of understanding that it can achieve. To do this, we have developed a theory of questions and their role in understanding and explanation.

Our approach raises several issues:

- Where do questions come from? What are the points in the understanding process where questions arise?

- What role do questions play in understanding and explanation? How do they affect understanding?

- How are questions indexed in memory such that they can get triggered when relevant input comes in?

- How do questions get transformed into new questions as new information comes in?

This paper is primarily about the first two questions, though we are addressing all four in our research. To contrast our approach with previous approaches to story understanding, let us consider a program such as FRUMP [DeJ79] as a question generation program. FRUMP had a database of *scripts* (also called frames or schemas) for different situations, such as terrorism and earthquakes. Each script contained a set of slots to be filled in when understanding a story about that kind of situation. For example, the earthquake script wanted to know the Richter scale reading, the number of people killed, and so on. The slots, therefore, represented the questions that the system asked every time it read about an earthquake. They also represented the limit of what the system could understand about earthquake stories. FRUMP would miss the point of a story about an earthquake in Pisa in which the Leaning Tower was destroyed, because it simply didn't have a slot for "famous monuments destroyed" in its earthquake script. In other words, it would never think of asking the question.

There are two ways out of this. We could, of course, add the missing slot to the other slots in the script, along with all the other slots we might need. Clearly it would be impossible to stuff all the required knowledge for all possible situations into a machine. We might compromise and stuff in a "lot" of knowledge as a start. But a machine that relied only on previously built-in knowledge would be able to understand just the situations that it was designed for. In order to be considered intelligent, we would want it to be able to deal with novel situations that it didn't already have the knowledge to deal with. In addition, all slots in all scripts are not equally interesting, so we would still have the problem of deciding which slots are interesting in a given situation. Most story understanders avoid this issue and pursue all of them with equal enthusiasm.

In our research, we have taken a different route. We have designed our system as a question generation program. The system asks questions as it processes a story, and then uses these questions to drive the understanding process. As a consequence, the system is interested in those facts that are relevant to the questions that it currently has. Thus the Richter scale reading of an earthquake would be interesting only if it was actually relevant to something it wanted to find out, and not simply because it was a slot in the earthquake script.

II. A taxonomy of questions

In order to understand where questions come from, as well as how they affect processing, we categorize questions into various types. The categories are defined in terms of the origin and functional role of questions in understanding. The taxonomy is based on informal data collected from several subjects. We will first present our taxonomy, and in the next section we will relate it to the explanation cycle that underlies the process of understanding.

Questions can be divided into five major categories:

- Explanation questions
- Elaboration questions
- Hypothesis verification and discrimination questions
- Reminding questions
- General interest questions

A. Explanation questions

Since constructing explanations is an important part of understanding, we would expect many questions to be concerned with explanation. Asking the right question is central to constructing the best explanation. An important class of questions, therefore, are *explanation questions* (or EQs). EQs focus our attention on a particular aspect of the situation, or allow us to view a situation in a particular way, with the intention of finding an explanation that might underlie it.

There are two major types of explanation questions. Since explanations are constructed to resolve contradictions or anomalies in the situation, EQs are often concerned with anomalies. For example, *Did the boy want the results of his actions?* is an *anomaly detection question* since thinking about this allows us to notice the anomaly in the first place. Given a characterization of an anomaly, we ask *anomaly resolution questions* to search for explanations of a particular type so that the situation isn't anomalous any more. For example, *Did the boy know he was going to die?* is an anomaly resolution question, since if he didn't this particular anomaly goes away.

The other kind of explanation questions seek stereotypical *explanation patterns* that might apply to the current situation. Explanation patterns, or XPs, are stock explanations that we have for various situations [Sch86]. For example, "Shiite religious fanatic does terrorism" is a standard XP many people have about the Middle East terrorism problem. We might think of them as the "scripts" of the explanation domain. When we see a situation for which we have a canned XP, we try to apply the XP to avoid detailed analysis of the situation from scratch. Explanation patterns are retrieved via explanation questions. For example, the question "Why would a Lebanese person perform a terrorist act?" has the religious fanatic explanation (and possibly others) indexed under it. The purpose of this question is to allow us to find these XPs.

B. Elaboration questions

Once we have retrieved a set of candidate explanation patterns, we try to apply them to resolve the anomaly. Often an XP cannot be applied directly, or is too general. In such situations, we might *elaborate* appropriate pieces of the explanation, or perhaps collect more information about the input. (There is also the possibility of *tweaking* the explanation, as in the SWALE program [Kas86], which we will not deal with here.)

For example, consider the blackmail incident that the car bombing story above tells us about. This provides an explanation for the boy's actions, but the explanation is incomplete. Some of the questions in our data were concerned with elaborating the explanation, such as *What could he want more than his own life?* and *Why do they choose kids for these missions?*. To answer these questions, we can either search memory for old episodes that might contain relevant information, or wait for further input. In the latter case, we call the question a *data collection question*, because it seeks to collect additional data pertaining to a given hypothesis.

C. Hypothesis verification and discrimination questions

Even after we construct (or are given) a detailed explanation, we may not know for certain that it is the right one. In fact, we typically have more than one competing hypothesis about what the best explanation is. The validity of a hypothesis depends on the assumptions that we made while constructing it. For example, although it is pretty easy to apply the "Shiite religious fanatic" XP in the car bombing example (before we are explicitly told that he is not a fanatic), the explanation rests on the assumptions that he is a Shiite Moslem and he is very zealous about his religion. These assumptions then become *hypothesis verification questions* (or HVQs) for the religious fanatic hypothesis: *Was he a Shiite Moslem?* and *Was he very zealous about his religion?*.

The role of HVQs is to verify or refute the hypothesis that they were generated from when answers to them are found. In case we have two or more competing hypotheses, they also help us to discriminate between the alternatives. Thus they represent what the understander is *interested* in finding out at any time for the purpose of understanding the story. However, unlike most story understanding programs, this notion of interestingness is *dynamic*. The boy's religion, for example, is interesting in this story because it is of relevance to the explanations begin constructed, and not because the "boy" frame has a "religion" slot that must always be filled.

D. Reminding questions

The fourth type of questions are *reminding questions*. The role of reminding in understanding is discussed in [Sch82] and [Sch86], and we will not pursue it here. Many questions are generated as a result of remindings based either on superficial features (e.g., *Why are they all named Mohammed?*), or on deeper explanatory similarities (e.g., *Are car bombers motivated like the Kamikaze?*). Reminding questions may suggest possible explanations stored with old episodes as candidate hypotheses for the current situation; they may also help us verify or refute hypotheses by providing supporting or opposing evidence from episodic memory. They also help us learn new categories by generalization over similar instances [Leb80] or over similar explanations.

E. General interest questions

Finally, we have questions already extant in memory before we begin to read the story. These questions are left over from our previous experiences. As we read, we remember these questions and think about them again in a new light. Certainly after reading the car bombing story, we expect to have several questions representing issues we were wondering about which weren't resolved by the story. For example, in this story it turns out that the boy was

blackmailed into going on the bombing mission by threatening his parents. This makes us think about the question *What are family relations like in Lebanon?*, which remains in memory after we have finished reading the story. To the extent that we are interested in this question, we will read stories about the social life in Lebanon, and we will relate other stories to this one. To cite another example, one of the students we read the story to repeatedly related the story to the IRA because he was interested in similar issues about Ireland.

Thus understanding is a process of question generation, and is in turn driven by these questions themselves. The traditional view of understanding is one of a process that takes a story as input and builds a representation of what it has understood. In contrast to this, we view understanding as a process that starts with questions in memory and, as a result of reading a story, answers some of them and generates a new set of questions to think about. Thus questions represent the dynamic "knowledge goals" of the understander.

III.   AQUA — A computer model

We have implemented a computer program called AQUA[1] which embodies our theory of questions and understanding. AQUA reads newspaper stories about terrorism and attempts to understand them by constructing causal and motivational explanations for the events in the stories. The explanations it constructs may be divided into four major levels. Each level corresponds to a set of explanation questions (EQs) that organize explanations at that level.

Action level: Explanations involving direct relationships between actions. For example, the question *Was the mission instrumental to another action that the boy wanted to perform?* is an EQ at this level.

Outcome level: Explanations involving direct benefits of actions for participants. For example, the question *Did the boy want the results of his actions?* is an EQ at this level.

Stereotype level: Explanations constructed from stereotypical explanation patterns (XPs). EQs at this level are *Why do teenagers commit suicide?* and *Why do Lebanese people perform terrorist acts?*.

Decision level: Ab initio reasoning about planning decisions. For example, if an action has a negative outcome for an agent who chose to perform the action (as opposed to being forced into it), we might ask the following questions:

- Did the agent know the outcome the action would have for him?
- Did the agent want that outcome (i.e., were we mistaken in assuming that the outcome was negative)?

- Was there another result of the action that the agent wanted, and did he want that result more than he wanted to avoid the negative result?

AQUA is part of an on-going project and is in the process of being developed. At present it reads the car bombing story mentioned above, but we are in the process of extending it to read other stories.

IV.   The explanation cycle

The processing cycle in AQUA consists of three interactive steps: **read, explain** and **generalize**[2]. AQUA starts with **read**ing some text and retrieving relevant memory structures to integrate new input into. This is guided by the questions that are currently in memory. These questions are generated during the **explain** step and indexed in memory to enable the **read** step to find them.

A.   The explain step

The **explain** step may be summarized as follows. Assume that AQUA has just read a piece of text.

Formulate EQ of appropriate type.

Retrieve XPs using EQ and general interest in certain types of explanations. For example, we might look for a social explanation for why a 16 year old Lebanese boy might want to commit suicide.

Apply XP to input:

If in applying the XP we detect an anomaly:

Characterize the anomaly.

Elaborate, using the anomaly characterization to focus the elaboration.

Explain the anomaly recursively, using the above characterization to guide the formulation of new EQs.

If the XP is applicable to the input:

Construct hypothesis by instantiating the explanation pattern.

Construct HVQs to help verify or refute the new hypothesis.

Index HVQs in memory to allow us to find them in the **read** step below.

If we can't apply the XP, try another one. If there are no more XPs, try a different EQ.

B.   The read step

At this point, AQUA has finished processing the newly read piece of text. It now continues reading the story, guided by the questions it has generated so far, as follows:

Read some text, focussing attention on interesting input as determined below.

---

[1] *Asking Questions and Understanding Answers.*

[2] The **generalize** step has not yet been implemented.

Retrieve extant questions indexed in memory that might be relevant. Use these questions as an interestingness measure to focus the **read** above.

Answer the questions retrieved in the previous step:

Answer HVQs by either confirming or refuting them.

Propagate back to the hypothesis that the question originated from.

Confirm/refute hypotheses. If the HVQs of a hypothesis are confirmed, confirm the hypothesis and refute its competitors. If any HVQ of a hypothesis is refuted, refute the corresponding hypothesis.

Explain the new input if necessary.

C. The generalize step

Since questions represent the difficulties encountered during understanding, they should also provide the focus for learning. A program such as AQUA should be able to:

Generalize the novel explanations encountered in story, using questions to focus generalization.

Index the generalizations back in memory, such that the original question which failed would now find it.

We have not yet implemented the **generalize** step in AQUA, but we are interested in this issue. Explanation-based learning algorithms [DeJ83, SCH86, SC82] involve the generalization of causal structures of explanations to form new generalized explanations while dropping the irrelevant details. However, programs such as GENESIS [MD85] that embody this idea perform undirected learning. We want to use questions as a mechanism to focus the learning process on the interesting or relevant aspects of the story.

V. Conclusions

We view story understanding as a process involving the generation of questions, which in turn drive further processing of the story. In this paper, we presented a taxonomy of the questions people ask while they read. We talked about the origin of these questions in the explanation cycle, and their role in understanding.

We are building a computer program to read and understand newspaper stories according to our theory. In contrast to the traditional view of understanding as a "story in, representations out" process, our program may be viewed as a "questions + story in, questions out" process. The paper presented our understanding algorithm as a three step integrated process: the **read** step, in which the program reads the text, focussed by the questions that are extant in its memory, the **explain** step, in which the program asks questions in order to construct explanations, and the **generalize** step, in which the program will generalize novel answers to its questions.

Other issues we are investigating include those of question transformation and judging the interestingness of different questions. We are also interested in using questions to direct learning.

Acknowledgements

References

[DeJ79] G. DeJong. *Skimming Stories in Real Time: An Experiment in Integrated Understanding.* Ph.D. thesis, Yale University, May 1979.

[DeJ83] G. DeJong. An Approach to Learning from Observation. In *Proceedings of the International Machine Learning Workshop,* 171–176, University of Illinois, Monticello, IL, June 1983.

[Kas86] A. Kass. *Modifying Explanations to Understand Stories. Proceedings of the Eighth Annual Conference of the Cognitive Science Society,* 691–696, Amherst, MA, August 1986.

[KLO86] A. Kass, D. Leake and C. Owens. *SWALE: A Program That Explains.* In [Sch86].

[Leb80] M. Lebowitz. *Generalization and Memory in an Integrated Understanding System.* Ph.D. thesis, Yale University, October 1980.

[Leh78] W.G. Lehnert. *The Process of Question Answering.* Lawrence Erlbaum Associates, Hillsdale, New Jersey, 1978.

[MD85] R. Mooney and G. DeJong. Learning schemata for natural language processing. In *Proceedings of the 9th IJCAI,* 681–687, Los Angeles, CA, August 1985.

[Sch82] R.C. Schank. *Dynamic memory: A theory of learning in computers and people.* Cambridge University Press, 1982.

[Sch86] R.C. Schank. *Explanation Patterns: Understanding Mechanically and Creatively.* Lawrence Erlbaum Associates, Hillsdale, New Jersey, 1986.

[SC82] Schank, R.C. and Collins, G.C. *Looking at Learning* In *Proceedings of the ECAI,* 10–16, Paris, France, 1982.

[SCH86] R.C. Schank, G. Collins, and L. Hunter. Transcending inductive category formation in learning. *Behavioral and Brain Sciences,* 9(4), 1986.