QUESTION ANSWERING IN THE CONTEXT OF CAUSAL MECHANISMS

Arthur C. Graesser, Darold Hemphill, and Lawrence E. Brainerd
Department of Psychology and the Institute for Intelligent Systems
Memphis State University
Memphis, TN 38152

Abstract

A model of human question answering (called QUEST) accounts for the answers that adults produce when they answer different categories of open-class questions (such as why, how, when, what-enabled, and what-are-the-consequences). This project investigated the answers that adults generate when events are queried in the context of biological, technological, and physical mechanisms. According to QUEST, an event sequence in a scientific mechanism is represented as a causal network of events and states; a teleological goal hierarchy may also be superimposed on the causal network in biological and technological domains, but not in physical systems (e.g., rainfall, earthquake).

When questions are answered, QUEST systematically operates on the causal networks and goal hierarchies that underlie a causal mechanism. Answers to <u>how</u> and <u>enablement</u> questions sample causal antecedents of the queried event in the causal network; <u>consequence</u> questions sample causal consequents. Answers to <u>when</u> questions sample antecedents to a greater extent than consequents even though events from both directions furnish sensible answers. Answers to <u>why</u> questions sample both causal antecedents in the causal network and superordinate goals from goal hierarchies that exist in technological and biological knowledge structures.

Graesser and Franklin (in press) have developed a model of human question answering called QUEST. This model accounts for the answers that are produced when individuals answer different categories of open-class questions: why, how, when, where, enablement, consequence, etc. Questions are answered in the context of different types of material, including stories, scripts, and expository texts on scientific mechanisms. The components of QUEST have foundations in some previous models that specify the Q/A processes which operate on structured databases representing world knowledge (Allen, 1987; Brachman, 1983; Graesser, Robertson, & Anderson, 1981; Graesser & Murachver, 1985; Lehnert, 1978; Lehnert, Dyer, Johnson, Yang, & Harley, 1983; McKoewn, 1985).

This research focuses on short scientific texts that depict causal event sequences in biological, technological, and physical mechanisms. For example, suppose that an individual reads the the following text about a nuclear power plant.

Event 1: Atoms are split.
Event 2: Heat energy is released.
Event 3: The water in the surrounding tank is heated.
Event 4: Steam drives a series of turbines.
Event 5: The turbines produce electricity.

After reading this text, individuals were probed with five different categories of questions:

WHY Why is water in the surrounding tank heated?

HOW How is water in the surrounding tank heated?
ENABLE - What enables water in the surrounding to be heated?
WHEN When is water in the surrounding tank heated?
CONS - What are the consequences of water in the surrounding tank being heated?

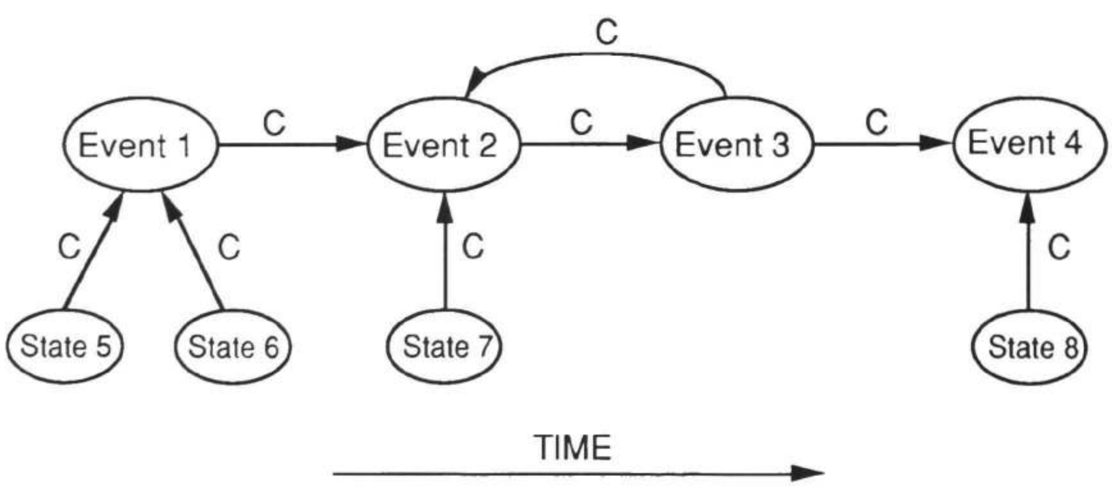
QUEST accounts for the answers that are produced when particular events are probed with a particular category of question. This project focuses on the "answer likelihood scores" of explicit events in the text. For example, what is the likelihood that Event 3 is produced when Event 2 is probed with a CONS question?

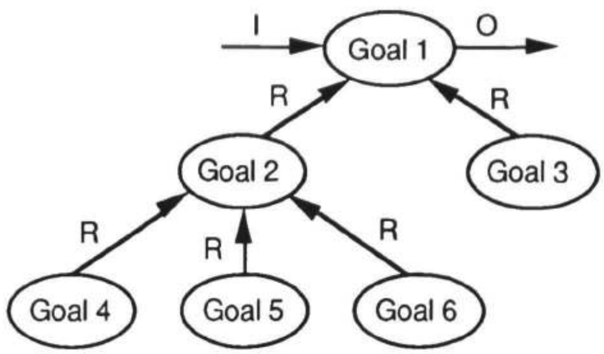
Representation of Knowledge

QUEST assumes that causal networks organize the events that are explicitly mentioned in event sequences. In the example causal network below, there is an event sequence that unfolds chronologically and causally (events 1, 2, 3, and 4). There may be some loops in the event sequence (as in events 2 and 3). A set of enabling states (states 5, 6, 7, and 8) are needed for the event sequence to unfold.

The Consequence (C) arcs in a causal network denote a weak sense of causality. If a forward C-arc connects nodes X and Y, then the connection must satisfy both temporal and causal criteria. Regarding temporality, node X must occur or exist prior to node Y. Regarding causality, X must furnish one or more of several causal relationships with Y. For example, a necessity relationship exists whenever a counterfactual test is satisfied: if X is negated or removed from the system, then Y will not occur. Other types of causal relationships involve sufficiency (X is sufficient for Y to occur), and operativity (X must be operating/existing when Y occurs). The causal analysis is similar to Trabasso's causal chain theory (Trabasso & van den Broek, 1985) and some theories in qualitative physics (Forbus, 1985).

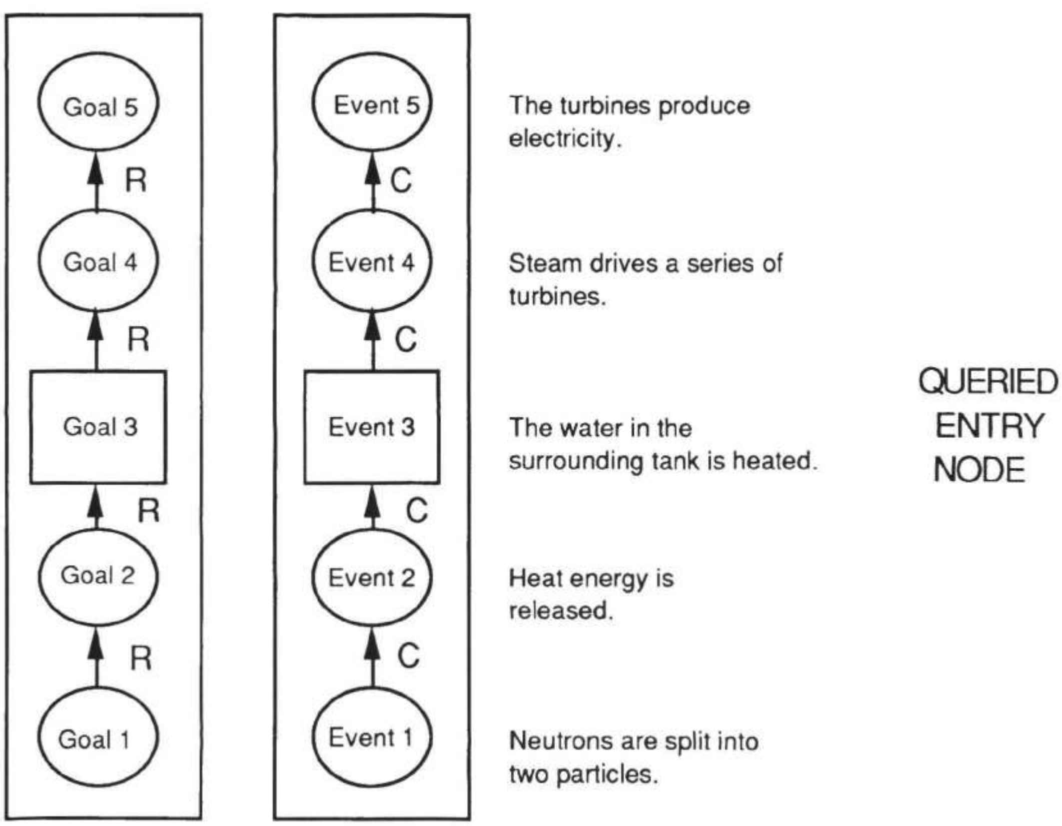
Our naive understanding of a scientific system may often be teleological. That is, events are organized according to a functional, purposeful, goal-oriented knowledge structure. For example, a nuclear power plant is designed to achieve the goal of producing electricity; the plan of achieving this goal is superimposed on the causal network of events and states. Technological systems are typically motivated by the goals and plans of animate agents. Biological systems are often interpreted teleologically. According to a naive view of evolution, for example, a species develops in order to acquire a superior capacity. Physical systems (e.g., rainfall, earthquakes) received a goal-oriented interpretation when the Greeks had myths about gods governing the natural events and processes.





Goal-oriented knowledge is represented as a hierarchy of goals that are related by Reason (R) arcs, as shown above. The most superordinate goal (goal 1) is the primary goal that the agent wants to achieve. The subordinate goals correspond to intermediate plans and low-level actions or processes. States/events in the world initiate the goals (signified by the I-arc); the goals may have outcomes (signified by the O-arc) that are positive or negative. Goal/plan hierarchies are very popular in theories of action and problem solving (Miller, Galanter, & Pribram, 1960; Graesser & Clark, 1985; Newell & Simon, 1972)

When there is a teleological understanding of a causal mechanism, a goal hierarchy is superimposed on the causal network. This is illustrated below in the context of the short text on nuclear power. For every event, there is a corresponding goal. Just as event 5 is the final event in the event chain, goal 5 is the most superordinate goal in the goal hierarchy. Just as event 1 is the initial event in the chain, goal 1 is the most subordinate goal in the goal hierarchy.



Question Answering Procedures

According to QUEST, each question category has an arc search procedure which specifies the legal path of arcs between (a) an "entry node" (the event being queried) and (b) a "good answer" node in the knowledge structure. For example, a question category might sample nodes that are causal antecedents to the entry node. Causal antecedent nodes are on paths that radiate from the entry node via backwards C-arcs in the causal network. When event 3 is probed with a HOW-question (How is water in the surrounding tank heated?), the antecedent nodes would be events 1 and 2, but not 4 and 5. The arc search procedures are specified for WHY, HOW, ENABLE, WHEN, and CONS questions in this paper.

Materials and Question Answering Protocols

The texts were 24 event sequences that were extracted from passages in the American Academic Encyclopedia. All texts had five events, as with the example on nuclear power. Eight passages were in the technological domain (e.g., television, nuclear power), 8 were in the biological domain (e.g., mitosis, photosynthesis), and 8 were in the physical science domain (e.g., earthquake, rainfall).

Question answering (Q/A) protocols were collected from college students at Memphis State University. Eight subjects provided Q/A protocols for each text, yielding 192 subjects altogether. Given that a text had 5 sentences and that there were 5 question categories (WHY, HOW, ENABLE, WHEN, and CONS), each subject answered 25 questions. The "answer likelihood" score for a particular answer was the proportion of subjects (out of 8) who gave that answer to a particular question.

Table 1 shows mean answer likelihood scores for those events that were causal antecedents versus causal consequents of the queried events. A causal antecedent occurred prior to the queried event (via a backwards C-arc or R-arc) whereas a causal consequent occurred subsequent to the queried event (via a forward C-arc or R-arc). The answer likelihood scores are segregated according to question category (averaging over the three knowledge domains). In addition, WHY-questions are segregated according to technological, biological, versus physical domains. These means are interpreted in the next section. In tests of statistical significance, the error term reflects the variability among items (i.e., a particular answer A to a question Q in a text T).

Table 1

Question Category	Antecedent Event	Consequent Event
Why?	.16	.17
How?	.21	.06
What enables?	.26	.08
When?	.36	.10
What are the consequences?	.08	.36
Why-Questions Segregated by Knowledge Domain		
Physical	.30	.10
Technological	.08	.20
Biological	.10	.20

Analysis of Five Question Categories

WHY-questions. The answers to WHY-questions were evenly divided between causal antecedents and causal consequents, according to Table 1. However, causal consequents were prevalent in technological and biological domains, but not in physical domains (see bottom of Table 1). In physical science domains, the nodes in causal networks are sampled by pursuing paths of backwards C-arcs from the entry node; goal hierarchies are not very prevalent. In technological and biological domains, goal hierarchies are more prevalent; the answers to why-questions sample superordinate goals via forward R-arcs (Graesser et al., 1981; Graesser & Clark, 1985; Graesser & Franklin, in press). Therefore, teleological interpretations are imposed on technological and biological domains, but rarely on physical science domains.

HOW-questions. Answers to HOW-questions included causal antecedents to the entry node in the causal network, as predicted by QUEST. The fact that causal consequents were occasionally sampled (answer likelihood = .06) can be attributed to occasional loops in the event chain and to answerers occasionally sampling the text in a haphazard manner. QUEST also predicts that some answers to HOW-questions specify the speed, force, intensity, and/or qualitative manner in which an event occurs (e.g., X occurred quickly, forcefully, in circles).

ENABLE-questions. Answers to these questions include causal antecedents to the entry node in the causal network, as predicted by QUEST. Thus, the answers to HOW and ENABLE questions are quite similar.

WHEN-questions. Strictly speaking, it is possible to relate a queried event temporally with either a causal antecedent (i.e., queried event occurred after the answer event) or with a causal consequent (i.e., queried event occurred before the answer event). However, there was a strong bias toward antecedent events in this study and in earlier research on narrative passages (Graesser & Murachver, 1985). The answers specified that a queried event occurs after its antecedent rather than before its consequent. According to QUEST, answers to WHEN-questions may also refer to higher-order, global event descriptions; the queried event would occur during the more global event.

CONS-questions. As predicted by QUEST, these answers included causal consequents of the entry node in the causal network (via paths of forward C-arcs). Answers to CONS-questions are essentially opposite of the answers to HOW, ENABLE, and WHEN questions.

Distance effects. Answer likelihood scores decreased as a function of the distance between the entry node and the answer node. A rather simple mathematical model provided a very good fit to the complete set of answer likelihood scores. The complete set includes 20 scores for each passage (5 queried events and 4 explicit answers per event), 24 passages, and 5 question categories (yielding 2400 scores altogether).

The mathematical model has three free parameters associated with each passage. Parameter $\underline{a}$ is the likelihood of pursuing a causal antecedent path whereas parameter $\underline{c}$ is the likelihood of pursuing a causal consequence path. Parameter $\underline{t}$ is the likelihood of traversing a single arc on a path whereas $\underline{n}$ is the number of arcs between the entry node and the answer node. The answer likelihood scores are closely predicted by $a*t^n$ for causal antecedents and $c*t^n$ for causal consequents. The overall mean value was .67 for $\underline{t}$, the distance dampening parameter. Regarding the causal antecedent parameter ($\underline{a}$), the mean value was substantially higher in those cells that are predicted as good answers by QUEST (.63) than in those cells that are not predicted to be good answers (.18). Regarding the causal consequent parameter (c), again the mean was much higher in good answer cells (.51) than in bad answer cells (.19).

To summarize, when distance effects are combined with the arc search procedures of QUEST, a very simple mathematical model provides an impressive fit to the answer likelihood scores. It should be emphasized once again, however, that an adequate account of answers to WHY-questions must assume the existence of (1) teleological goal hierarchies being superimposed on causal networks and (2) systematic differences among technological, biological, and physical science mechanisms.

References

Allen, J. (1987). Natural language understanding. Reading, MA: Benjamin/Cummings.

Brachman, R. J. (1983). What IS-A is and isn't: An analysis of taxonomic links in semantic networks. Computer, 16, 30-36.

Forbus, K. D. (1985). Qualitative process theory. In D. G. Bobrow (Ed), Qualitative reasoning about physical systems. Cambridge, MA: MIT Press.

Graesser, A.C., & Clark, L.F. (1985). Structures and procedures of implicit knowledge. Norwood, NJ: Ablex.

Graesser, A.C., & Franklin, S.P. (in press). QUEST: A cognitive model of question answering. Questioning Exchange.

Graesser, A.C., & Murachver, T. (1985). Symbolic procedures of question answering. In A.C. Graesser & J.B. Black (Eds.), The psychology of questions. Hillsdale, NJ: Erlbaum.

Graesser, A.C., Robertson, S.P., & Anderson, P.A. (1981). Incorporating inferences in narrative representations: A study of how and why. Cognitive Psychology, 13, 1-26.

Lehnert, W. G. (1978) The process of question answering. Hillsdale, NJ: Erlbaum.

Lehnert, W. G., Dyer, M. G., Johnson, P. N., Yang, C. J., & Harley, S. (1983). BORIS: An in-depth understander of narratives. Artificial Intelligence, 20, 15-62.

McKoewn, K. R. (1985). Discourse strategies for generating natural language text. Artificial Intelligence, 27, 1-41.

Miller, G. A., Galanter, E. & Pribram, K. H. (1960). Plans and the structure of behavior. New York: Holt, Rinehart, & Winston.

Newell, A., & Simon, H. A. (1972). Human problem solving. Englewood Cliffs, NJ: Prentice-Hall, Inc.

Trabasso, T., & van den Broek, P. (1985). Causal thinking and the representation of narrative events. Journal of Memory and Language, 24, 612-630.