Knowledge Goals: A Theory of Interestingness*

Ashwin Ram

Georgia Institute of Technology
School of Information and Computer Science
Atlanta, Georgia 30332-0280
(404) 853-9372
E-mail: ashwin@ics.gatech.edu

Abstract

Combinatorial explosion of inferences has always been one of the classic problems in AI. Resources are limited, and inferences potentially infinite; a reasoner needs to be able to determine which inferences are useful to draw from a given piece of text. But unless one considers the goals of the reasoner, it is very difficult to give a principled definition of what it means for an inference to be "useful."

This paper presents a theory of inference control based on the notion of interestingness. We introduce *knowledge goals*, the goals of a reasoner to acquire some piece of knowledge required for a reasoning task, as the focussing criteria for inference control. We argue that knowledge goals correspond to the *interests* of the reasoner, and present a theory of interestingness that is functionally motivated by consideration of the needs of the reasoner. Although we use story understanding as the reasoning task, many of the arguments carry over to other cognitive tasks as well.

1 Cognitive motivations: Knowledge goals as a basis for interestingness

When we compare the way people read newspaper stories with how computer programs typically read them, we notice the following differences:

Subjectivity: People are biased. They interpret stories in a manner that suits them. They jump to conclusions. Computer programs, on the other hand, are usually designed to read stories in an objective manner, and to extract the "correct" or "true" interpretation of a story to the extent that they can.

Variable depth parsing: People don't read everything in great detail. They concentrate on details that they find relevant or interesting, and skim over the rest. In contrast, computer programs are designed to attend to every aspect of a story that is within the scope of their knowledge structures. Consequently, they either process the entire story in great depth, or else they skim everything in the story. They can not decide which aspects to process in detail and which ones to ignore.

Learning and change: People change as they read. They never read the same story twice in the same way. They notice different things the second time around, or they simply get bored. After reading a story, they interpret other similar stories differently. Computer programs, in contrast, are not adaptive; they always read a given story the same way.

What makes people different from computer programs? What is the missing element that our theories don't yet account for? The answer is simple: *People read newspaper stories for a reason: to learn more about what they are interested in.* Computers, on the other hand, don't. In fact, computers don't even have interests; there is nothing in particular that they are trying to find out when they read. If a computer program is to be a model of story understanding, it should also read for a "purpose."

Of course, people have several goals that do not make sense to attribute to computers. One might read a restaurant guide in order to satisfy hunger or entertainment goals, or to find a good place to go for a business lunch. Computers do not get hungry, and computers do not have business lunches.

However, these *physiological* and *social goals* give rise to several *intellectual* or *cognitive goals*. A goal to satisfy hunger gives rise to goals to find information: the name of a restaurant which serves the desired type of food, how expensive the restaurant is, the location of the restaurant, etc. These are goals to acquire information or knowledge, and are called *knowledge goals*. These goals can be held by computers too; a computer might "want" to find out the location of a restaurant, and read a guide in order to do so in the same way as a person might. While such a goal would not arise out of hunger in the case of the computer, it might well arise out of the "goal" to learn more about restaurants.

In other words, knowledge goals also arise from

*The research described was conducted while the author was at Yale University, and supported by the Defense Advanced Research Projects Agency and the Office of Naval Research under contract N00014-85-K-0108, and by the Air Force Office of Scientific Research under contracts F49620-88-C-0058 and AFOSR-85-0343.

the desire to learn, to pursue one's intellectual interests, to improve one's model of the world. These goals can be viewed as *questions* about the domain of interest. To be interested in terrorism, for example, is to have a lot of questions about various aspects of terrorism, and to think about these questions in the context of input data about terrorism, such as newspaper stories about terrorist incidents. The point of reading these stories is to answer one's questions, as well as to reveal flaws or gaps in one's model of terrorism in order to try to improve this model. These gaps give rise to new questions which in turn stimulate further interest in terrorism. Both computers and people can be "interested" in terrorism in this sense.

In contrast with people, therefore, a computer has only one underlying goal: to learn and improve its world model.[1] However, this (and, in the case of people, other physical and social goals) gives rise to knowledge goals that then drive the understanding process.

2 Computational motivations: What are the knowledge goals of an understanding program?

Understanding, then, can be viewed as the pursuit of one's interests or questions. However, it would defeat the purpose to build a "question-asking" or "interest-pursuing" program per se. Instead, these questions and interests should arise naturally as *cognitive goals* of the program during various stages of the reasoning process. This means that the program should ask a question only when it has a need to acquire that piece of knowledge. For example, in the case of a story undertanding system, a knowledge goal should be formulated only when the system needs to know the answer for the purposes of understanding the story. In other words, knowledge goals should be *functionally useful* to the overall goals of the system.

The theory of knowledge goals presented in this paper depends on a theory of *understanding tasks*, the basic tasks of an understander. In addition to parser-level tasks such as noun group connection, pronoun reference, etc., these tasks include the integration of facts with what the understander already knows, the detection of anomalies in the text which identify flaws or gaps in the understander's model of the domain, the formulation of explanations to resolve those anomalies, the confirmation and refutation of potential explanations, the learning of new explanations for use in understanding future situations. These are the basic tasks that an understander needs to be able to perform.

In order to carry out these tasks, the understander needs to integrate the text, which is often ambiguous, elliptic and vague, with its world knowledge, which is often incomplete. In formulating an explanation, for example, the understander may need to know more about the situation than is explicitly stated before it can decide which is the best explanation. However, it is impossible to anticipate when a particular piece of knowledge will be available to the understander, since the real world (in the case of a story understanding program, the story) will not always provide exactly that piece of knowledge at exactly the time that the understander requires it. *Thus the understander must be able to suspend questions in memory, and reactivate them at the right time when the information it needs becomes available.* In other words, the understander must be able to remember what it needs to know, and why.

Furthermore, the system's understanding of any real world domain can never be quite complete. Conventional script, frame or schema-based theories assume that understanding means finding an appropriate script, frame or schema in memory and fitting it to the story. Schemas in memory are assumed to be "correct;" if an applicable schema is found, the story is understood. However, this model is inadequate since an understander's memory is always incomplete. Knowledge structures often have gaps in them, especially in poorly understood domains. These gaps correspond to what the understander has not yet understood about the domain. Even if a schema appears to be correct, novel experiences or stories may reveal flaws in the schema or a mismatch with the real world. Furthermore, the schema may not be indexed correctly in memory.

Understanding tasks, therefore, generate information subgoals or questions, representing what the understander needs to know in order to carry out the current task, be it explanation, learning, or any other cognitive task. These questions constitute the specific knowledge goals of the system, and are used to focus the understanding process.

Our theory of knowledge goals is motivated by these functional considerations, and corresponds well with a theory of interestingness motivated by the above cognitive considerations. The theory has been implemented in a computer program called AQUA (Asking Questions and Understanding Answers), which learns about terrorism by reading newspaper stories about unusual terrorist incidents in the Middle East [Ram, 1987; Schank and Ram, 1988; Ram, 1989]. AQUA uses its knowledge goals to direct the understanding process. We will illustrate our ideas with examples taken from this program.

3 A taxonomy of knowledge goals

Knowledge goals can be characterized according to the type of understanding task that they arise from.

Text goals: Knowledge goals of a text analysis program, arising from text-level tasks. These are the questions that arise from basic syntactic and semantic analysis that needs to be done on the input

[1] Since computers will eventually be expected to interact with the physical world (e.g., robots) and the social world (e.g., employees), they will also be expected to have some of the physical or social goals that we currently attribute only to people.

text, such as noun group attachment or pronoun reference.

Memory goals: Knowledge goals of a dynamic memory program, arising from memory-level tasks. A dynamic memory must be able to notice similarities, match incoming concepts to stereotypes in memory, form generalizations, and so on.

Explanation goals: Goals of an explainer that arise from explanation-level tasks, including the detection and resolution of anomalies, and the building of 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.

Relevance goals: Goals of any intelligent system in the real world, concerning the identification of aspects of the current situation that are "interesting" or relevant to its own goals.

AQUA is an implementation of a integrated theory of story understanding, memory and learning, called *question-driven understanding*, which addresses the above issues. In addition to their theoretical role in our model of inference control and interestingness, knowledge goals have also played an implementational role in our research by providing a uniform mechanism for the integration of various cognitive processes. For example, knowledge goals arising from, say, memory tasks are indexed in memory and used in the same way as knowledge goals arising from explanation tasks. A knowledge goal generated from one task may be suspended, and satisfied opportunistically during the pursuit of some other task at a later stage or even during the processing of a different story. Implementational details may be found in [Ram, 1989].

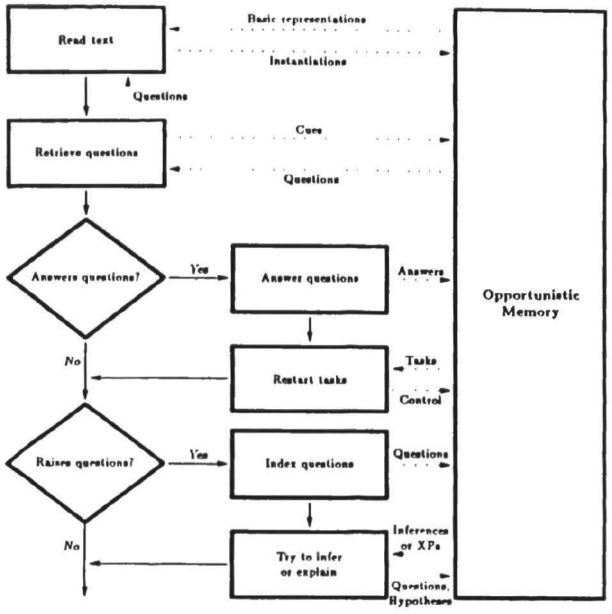

Figure 1: Control structure: The understanding cycle. A fact is interesting if it satisfies a knowledge goal pending in memory, or if it gives rise to new knowledge goals. Uninteresting facts pass vertically down with minimal processing; interesting facts cause suspended understanding tasks to be restarted, or new tasks to be created. New tasks can give rise to new knowledge goals, which are suspended along with the tasks if answers are not yet known and cannot be inferred.

4 Using knowledge goals to guide processing

A program that uses knowledge goals to guide understanding is an improvement over one that processes everything in equal detail, i.e., one that is completely text-driven. An understander that is completely text-driven would process everything in detail in the hope that it might turn out to be relevant. To avoid this, the understander should draw only those inferences which would help it find out what it needs to know. In other words, the understander should use its knowledge goals to focus its attention on the interesting aspects of the story, where "interesting" can be defined as "relating to something the understander wants to find out about."

Why would an understander need to find something out in the first place? Ultimately, the point of reading is to learn more about the world. Questions arise when reading a story reveals gaps or inconsistencies in the world model. It is useful to focus attention on such questions because they arise from a "need to learn." For example, questions arising from anomalous facts are more useful than those arising from routine stereotypical facts, since in the former case the understander may learn something new about the world.

There are two basic ways in which a fact can turn out to be worth processing:

Top-down: A fact that answers a question is worth focussing on since it helps to achieve a knowledge goal of the understander, which in turn allows the understander to continue the reasoning task that was awaiting the answer.

Bottom-up: A fact that raises new questions is worth focussing if the questions arise from a gap or inconsistency in the understander's knowledge base, since the understander may be able to improve its knowledge base by learning something new about the world.

These correspond to the two diamonds in figure 1. These diamonds attempt to determine which facts the understander should focus on. To improve on this even further, the understander needs a way of determining which knowledge goals worth pursuing and which ones are not. Not all questions are equally important, nor are all answers equally valuable. The understander needs to be able to determine the priorities of its knowledge goals, depending

on how likely the understander is to learn something by thinking about these knowledge goals. These decisions are made using a set of heuristics that will be described below.

The decision to focus attention corresponds closely with the notion of "interestingness." When an understander focuses on a particular fact and processes it in greater detail, it can be said to be "interested" in that fact.[2] For this reason, focus of attention heuristics can also be thought of as *interestingness heuristics*. These heuristics provide a functional definition of "interestingness" as a criterion for focussing attention: *Interestingness is a guess at what one thinks one might learn from paying attention to a fact or a question.* The guess must be made without processing the fact or question in detail, because otherwise the purpose of focussing attention to control inferences would be defeated. Thus the interestingness heuristics described below are indeed *heuristics* rather than precise measures of the value of thinking about a fact or a question.

5 Interestingness heuristics

In order to use questions to control inferences, an understander must be able to determine the interestingness of questions based on their content, as well as the task that they arose from. It must also be able to identify facts in the story that are interesting by virtue of being relevant to questions that the understander is interested in.

There are two types of heuristics for determining interestingness:[3]

Content-based: The interestingness of some input depends on its content or domain (more specifically, on the relationship between its content and the system's goals). In other words, some things are more interesting than others, depending on their relationship to the system's goals. For example, if one is intending to fly KLM in the near future, a story about a KLM flight being hijacked would be very interesting even if it were a stereotypical hijacking story. The understander would try to draw those inferences that were relevant to its goal of flying KLM. Similarly, stories about people one personally knows are more interesting than stories about strangers.

These heuristics use the content of the fact or question to determine its interestingness. The issue here is, which particular facts should the understander focus on? Which particular facts does the understander need to learn about?

Structure- or Configuration-based: Some kinds of situations are more interesting than others. For example, expectation failures are interesting, regardless of the content of the particular expectation that failed. These heuristics use the structure of the knowledge to determine interestingness. The issue here is, can the structure of the situation be used to determine which aspects of the situation are worth focussing on? Which configurations of knowledge structures signal gaps that the understander needs to learn about?

Of course, in any particular situation, the two conditions need to be combined in order to determine the overall interestingness of the input. For example, an expectation failure that relates to a goal of the system would have a higher priority than one that does not, and so the system would be more interested in the former.

Both types of heuristics identify situations in which the understander might be able to learn something useful. In the first case, the understander might learn more about a person it knows, or it might learn a new way to achieve a goal that it has. In the second case, the understander might be able to update its world model by identifying gaps in its model.

An example of the first type of heuristic is the principle of goal identification used by the POLITICS program [Carbonell, 1979]:

> If the understander of an event identifies with the goals of one of the actors, he will focus attention on inferences that lead to the fulfillment of these goals.

This is a content-based heuristic since it relies on the actual type of goal, not merely on the fact that there is a goal being pursued. A configuration-based heuristic, on the other hand, relies on particular relationships between concepts in memory, not specifically on what those concepts are. For example, POLITICS used the following configuration-based heuristic to focus its attention:

> Objective/Means distinction: If there are two or more actions in an event, and some actions are instrumental to stated or implicit objectives of one of the actors, the understander should focus attention on the objectives and non-instrumental actions.

In other words, an instrumental action is less likely to be significant than a larger action that it is part of, regardless of what the particular actions are. The reason this is a good strategy for an understander, according to the definition of interestingness proposed earlier, is that the understander is more likely to learn something by thinking about the larger action than it is by thinking about the instrumental action.

AQUA uses several heuristics to judge interestingness. These heuristics can be categorized according to the type of understanding goals that they pertain to. Let us start with relevance goals.

[2]Since interestingness depends on one's goals, the heuristics presented here do not cover interests that arise from goals that lie outside the scope of the basic understanding and learning tasks that AQUA performs. For example, a parent would be interested in the report card of his child. Since AQUA's goals do not include caring for children, it would not have any reason to be interested in a report card, unless the report card was anomalous with respect to AQUA's beliefs.

[3]This is orthogonal to the top-down/bottom-up distinction made earlier.

5.1 Interestingness from relevance goals

Interestingness arising from personal relevance usually falls into the class of content-based interestingness, since particular goals, people, locations, etc. are identified as being interesting to the understander. Questions and facts involving these goals, people, locations, etc. are worth pursuing since the understander might learn something relevant to it by doing so.

5.1.1 What could be relevant to a program? In order for something to be personally relevant to a program, the program must have a personality in the real world. There must be goals it wants to achieve, people it knows or has heard of, places it knows about or has grown up in, and so on. Stories relevant to this personality are interesting even if they do not involve anomalies or novel explanations.

For example, two of the focus of attention criteria used by POLITICS fall into this category: "goal identification" and "interest in VIP activities." Since POLITICS had a political ideology, it could be said to have a personality in the sense used here. Thus it could focus its attention on those aspects of a situation that were relevant to its goals.

Since AQUA does not have any real experiences outside of reading stories, its "personality" consists of its knowledge goals, i.e., the questions that it is interested in finding answers to. The people, institutions, objects and locations that it is interested in learning about are the people, institutions, objects and locations that are involved in these questions. The goals that are interesting to AQUA are goals of characters in these stories that it has questions about.

One could say that AQUA has "adopted" the goals that it has questions about, in the sense that it is interested in stories (or aspects of stories) about such goals as if they were its own (an example follows later). An alternative approach to the problem of where a program might get its goals is to "give" goals to the program, for example, by programming particular ideologies into the program as in POLITICS. In AQUA, this would be analogous to tagging particular goals or people as being "personally relevant" or "interesting" to the program.

In contrast, the approach used in AQUA is to allow the program to evolve its own set of interests that are functional to the purpose of the program (learning about terrorism), by letting the questions that arise from this purpose *be* its goals. These interests can be used to focus attention on those aspects of the story that would help it achieve its purpose.

Let us now discuss interestingness heuristics based on this notion of personal relevance.

5.1.2 Goal relevance. This is similar to the "goal identification" criterion used by POLITICS.

H-1: Goal relevance
A fact that could be instrumental to or could hinder a goal of the understander is more interesting than one that has no relevance to the understander's goals. A fact that directly matches or conflicts with a goal of the understander is very interesting.

5.1.3
Vicarious goals. Vicarious interests arise from goal configurations that are similar to one's own, or to those one is likely to have at some point. Stories about crisis goals or sudden goal changes, for example, are usually interesting from a vicarious point of view since the understander is likely to experience similar crisis goals or goal changes.

These heuristics are configuration-based, as opposed to content-based, since all goal changes are inherently interesting regardless of the particular goals involved. However, goal changes involving goals of personal relevance to the understander would be more interesting than those involving other goals.

H-2: Vicarious goal change
An action that drastically changes the goals of the planner or actor of the action is interesting.

H-3: Vicarious crisis goal
An action that initiates a crisis goal of the planner or actor of the action is interesting.

An example of a vicarious crisis goal is the self-preservation goal arising from acts of violence. Schank calls violence an "absolute interest" [Schank, 1979] since people universally seem to be interested in violence. This follows from the application of heuristic H-3 to the thematic goal of self-preservation which is universal among people.

5.1.4 Actor relevance. There are four degrees to which a particular person or institution can be relevant to an understander, each of which produces more interest than the previous one:

H-4: Actor relevance
Stories involving people that are completely unknown are the least interesting. More interesting than these are stories involving people who the understander has heard of. Still more interesting are stories involving people who are famous, such as celebrities. Stories involving people who the understander personally knows are the most interesting.

This heuristic is similar to the "actor relevance" and "interest in VIP activities" criteria of POLITICS. AQUA uses similar heuristics for object and location relevance, omitted here due to space limitations.

5.2 Interestingness from explanation goals

So far, we have seen interestingness heuristics that tried to identify situations involving goals, people, objects or places that are personally relevant to the understander. Such situations provide an opportunity to learn something of personal relevance, e.g.,

to learn something more about a person the understander is interested in, or to learn a new plan to achieve a goal that the understander has. There is another class of heuristics which identify potential gaps in the understander's knowledge, in order to determine what the understander might learn from processing the given situation. Situations that are interesting according to these heuristics are those that allow the understander to improve its world model. These heuristics are based on explanation and memory goals of the understander.[4] Let us start with explanation goals.

5.2.1 Anomaly detection. Anomalies arise when incoming facts do not fit in with what the understander expected to see. Anomalies are interesting because there is a possibility that the world model that underlies the failed expectations is incorrect, which signals a need to learn more about the domain.

H-5: Anomaly detection
All anomalies are interesting, and therefore all knowledge goals that arise during the anomaly detection process are interesting. These goals are always pursued in an attempt to form explanations to resolve the anomalies.

Knowledge goals that seek explanations in order to resolve anomalies are less interesting if explanations are easily available. If no explanation is found, there is a gap in memory corresponding to an unexplainable anomaly. A knowledge goal that seeks to fill in this gap is interesting for an understander that is trying to learn about this domain.

H-6: Explanation availability
An explanation retrieval goal that fails to find an explanation in memory is more interesting than one where an explanation can be easily found.

Thus stories that have standard explanations are less interesting than novel and unusual stories.

Of course, this and any other heuristic might be overruled if other interestingness heuristics came into play. Heuristics like these encode rules that represent guesses at interestingness, *other things being equal.* Counter-examples can easily be found by playing up interestingness factors represented by other heuristics. The final interestingness measure is a combination of all these factors, not any single factor taken by itself. For example, a stereotypical explanation for the motivations of a person who one knows personally might be more interesting than an unusual explanation for the motivations of a person one doesn't know. This also depends on the goals of the system, since a reasoner might be particularly interested in the motivations of a particular person for some other reason.

The interestingness of an explanation goal also depends on the kind of explanation that the goal is seeking. Knowledge goals seeking explanations for the motivations of individuals arise from *human interest stories*, and are usually more interesting, from the human interest point of view, than those that seek explanations for the motivations of institutions.

H-7: Human interest
Knowledge goals involving motivations of individuals are more interesting than those involving motivations of institutions.

Note that heuristic H-7 judges interestingness from a human interest point of view. Human interest stories focus on the goals, motivations and emotions of particular individuals. From the point of view of politics or counter-planning, however, the motivations of the institution might be more interesting than those of an individual, unless the individual was an important political figure in that institution.

When searching for explanations for actions, instrumental actions are less interesting than the actions that they are instrumental to. This is similar to the objective/means criterion of POLITICS that was described earlier.

H-8: Instrumentality
If an action is instrumental to or part of another action, the former action is less interesting than the latter. If an action has more than one action instrumental to it, or a MOP has more than one scene, the most interesting of the instrumental actions is that which is the goal scene of the MOP (or the "maincons" of a script).

Similarly, habitual or commonly performed actions are uninteresting by the following heuristic:

H-9: Thematic or stereotypical action
If the actor, or a group that the actor belongs to, is known to perform such actions or select such plans in service of a known thematic goal, the action is not interesting. Thus common plans for goals and routine thematic or occupation-related actions are uninteresting.

These heuristics are used to focus the understander's attention on the most interesting actions. Other actions are "explained away" by building simple explanations. For example, AQUA does not explain stereotypical actions in detail. The explanation it builds is simply "Because the actor often performs such actions in service of his goals," unless it has been unable to explain such actions in the past in which case it would have a pending question which would make this action interesting. Such heuristics allow AQUA to spend its time processing the more interesting aspects of the story. If these heuristics are absent, AQUA will still process the interesting aspects of the story, of course, but it will spend a lot more time processing uninteresting details as well.

5.2.2 Hypothesis formation. When a possible explanation is found, it is applied to the anomalous situation in order to construct a hypothesis that

[4] Since AQUA does not perform any text-level learning, text-level goals are not interesting to AQUA according to the learning criterion for interestingness.

might explain the anomaly. If a stereotypical explanation is available that applies easily and directly, the story conforms to the explanation that the understander already knows about, and is therefore not very interesting. On the other hand, if existing explanations do not apply to the situation, the story is novel and therefore interesting.

H-10: Hypothesis formation
If an available explanation applies easily and directly to the story, the story is not very interesting. More interesting is the case when a known explanation applies but leaves gaps which need to be filled in before the hypothesis is verified. The most interesting story is one in which known explanations do not fit the situation and need to be modified.

This heuristic follows from the claim that interestingness is a measure of what the understander might learn from processing the situation. Stories that identify gaps in the understander's memory are more interesting than those that fit into stereotypical molds that the understander already knows about, since the understander is unlikely to learn anything from processing the latter kind of stories.

5.2.3 Hypothesis verification. The final step in the explanation process is the verification of hypotheses. Facts in the story that are relevant to existing hypotheses are more interesting than facts that have no bearing on hypotheses currently in memory:

H-11: Hypothesis verification
An input fact is interesting if it helps to verify or refute a hypothesis that might explain an anomaly.

It is worth noting that these heuristics are dynamic and therefore a considerable improvement over static heuristics that select interesting features or facts on some arbitrary basis. For example, the color of an agent's hair is usually irrelevant in most stories. However, if this feature were statically marked as being uninteresting, an understander would be unable to correctly process a story in which this feature turned out to be interesting for some unforeseen reason.

5.3 Interestingness from memory goals

For an explanation-based program such as AQUA, explanation goals are more interesting than memory or text-level goals. However, memory-level tasks also give rise to heuristics for interestingness which, as before, are based on trying to identify gaps in memory which give the understander an opportunity to learn. For example, the basic learning mechanism in a program such as IPP [Lebowitz, 1980] is that of similarity-based generalization, a process that builds categories in memory by noticing similarities between instances or sub-categories and building generalizations based on these similarities. This theory of learning suggests the following heuristics:

H-12: Uniqueness
A category with a unique example or specialization is more interesting than one with several examples. If two categories have unique examples, the category higher in the type hierarchy is more interesting than the one that is lower down.

H-13: Symmetry
A category which lacks a symmetric category is more interesting than one for symmetric categories are known.

The symmetry can be along any of the dimensions used for similarity-based generalization in memory. For example, if the understander builds categories of occupations based on the gender of the actor, occupations in which both male and female actors are seen would be less interesting than those in which only males or only females are seen. This assumes, of course, that these categories play some functional role in achieving the overall goals of the understander, otherwise there would be no principled reason for either building the categories or judging their interestingness.

6 Computing interestingness by combining heuristics

We have presented a set of content- and structure-based heuristics for judging interestingness based on different types of understander goals. This set is not exhaustive, of course, but it illustrates the type of heuristics that an understander would use to determine interestingness and focus its attention. This allows the understanding process to be sensitive to the knowledge goals of the system.

The final measure of interestingness is derived by combining the recommendations of all the applicable heuristics. This is used to judge the interestingness of the system's knowledge goals, as well as the interestingness of facts that might be relevant to these knowledge goals.

H-14: To determine the interestingness of a fact or a knowledge goal, apply all the interestingness heuristics to the fact or knowledge goal and combine the interestingness recommendations of each heuristic.

There is a potential problem here since the heuristics given above don't recommend specific interestingness values. For example, is an anomaly involving an uninteresting goal more interesting than a stereotypical way of achieving a highly interesting goal? This problem has not yet been addressed. The current implementation of AQUA pursues every knowledge goal that is judged to be interesting by one or more heuristics.

7 Examples

In conclusion, let us illustrate the above interestingness heuristics by using them to determine the interestingness of some example stories from the terrorism domain. Consider the following story (New York Times, April 14, 1985):

S-1: Boy Says Lebanese Recruited Him as Car Bomber.

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 [by the Islamic Jihad] through another means: blackmail.

We can use the above heuristics to judge the interestingness of this story.

H-8 **Instrumentality:** The suicide bombing is not instrumental to a known larger plan.

H-3 **Vicarious crisis goal:** The boy's action affects his `preserve-life` goal.

H-6, H-10 **XP availability and applicability:** A stereotypical XP is available (religious fanatic) but inapplicable. Another stereotypical XP (blackmail) is applied in a novel context.

H-11 **Hypothesis verification:** The religious fanatic hypothesis is refuted. The blackmail explanation applies, but raises new questions such as "What could the boy want more than his own life?" which must be answered before the hypothesis is completely filled out.

H-7 **Human interest:** The explanation discusses personal motivation.

H-12 **Uniqueness:** The blackmail explanation has never been applied to a suicide bombing story before.

H-4 **Actor relevance:** The actor and planner are unknown to the understander.
Object relevance: The objects involved are unknown to the understander.

H-1 **Goal relevance:** No personal goals are achieved or violated.
Location relevance: Lebanon is not personally relevant to the understander.

This story, although not personally relevant, is interesting from the point of view of human interest. The story discusses novel explanations for the motivations behind the actions involving the violation of a shared thematic goal, `preserve-life`. Suppose AQUA has read several religious fanaticism stories, but has not encountered any coercion stories so far. After reading the above story, AQUA will be left with several questions, including:

- What did Mohammed value more than his own life?
- Why did the Islamic Jihad plan this mission?
- Why did the Islamic Jihad choose a teenager for the mission?

These questions are represented as knowledge goals in AQUA's memory. On the basis of these knowledge goals, AQUA will now be interested in stories involving the people or institutions it has questions about, such as:

- Another story about Mohammed Burro (relevance to known person)
- Another story about the Islamic Jihad (relevance to known institution)

AQUA will also be interested in stories involving the newly learned blackmail explanation, such as:

- Another story about someone being blackmailed into a suicide bombing mission (relevance to novel explanation)

AQUA will also be interested in a story involving goals or goal priorities that it has questions about, such as:

- Another story about someone valuing something over their own life (relevance to goal)

In contrast, consider the following more stereotypical story:

S-2: Suicide bomber strikes U.S. embassy in Beirut.

A teenage girl exploded a car bomb at the U.S. embassy in Beirut today, killing herself and causing a number of casualties, security sources said. A statement by an unidentified terrorist group claimed responsibility for the attack, adding that the girl was a martyr for the cause of Islam.

This story is relatively uninteresting. It is interesting only to the extent that it is about the U.S. embassy (personal relevance of object to an American understander), and that it discusses the motivations behind the violation of a vicarious crisis goal (heuristic H-3). However, since these motivations have a standard explanation that has been seen many times before (religious fanaticism), this story is not interesting even from the point of view of explanation (H-6, H-10).

8 Conclusion: Interest-producing conditions

We define interestingness as a criterion for inference control. Since the understander needs to focus its attention on those inferences likely to help it achieve its overall goals, it must devote its resources pursuing inferences that are most likely to be useful towards achieving these goals. Thus interestingness is a heuristic measure of the relevance of the input to the understander's knowledge goals. Since the point of satisfying knowledge goals is to improve one's understanding of the domain, interestingness can also be thought of as a measure of the likelihood of learning something from the story if one processes it in detail.

Interestingness is neither inherent in the information nor in the system, but rather arises from the interaction between the two. It arises from the interaction between the stimulus and the goals of the system. A system with no goals would have no reason to find any input more interesting than any other, nor would any particular piece of information be universally interesting for all systems unless they shared the same goals.

This is a functional approach to the problem of interestingness [Hidi and Baird, 1986; Schank, 1979] from the perspective of the theory of question-driven understanding. A similar approach can be used for systems performing other cognitive tasks, such as planning, since these systems would also need to focus their attention on inferences that were relevant to goals arising from their tasks.

In AQUA, interest in a concept is triggered by its likely relevance to questions or knowledge goals, and continuing interest is determined by its continuing significance to these goals. This is related to the "goal satisfaction principle" of [Hayes-Roth and Lesser, 1976], which states that more processing should be given to knowledge sources whose responses are most likely to satisfy processing goals, and to the "relevance principle" of [Sperber and Wilson, 1986], which states that humans pay attention only to information that seems relevant to them. These principles make sense because cognitive processes are geared to achieving a large cognitive effect for a small effort. To achieve this, the understander must focus its attention on what seems to it to be the most relevant information available [Sperber and Wilson, 1986].

Once the interestingness of a question or piece of input has been determined, AQUA uses it to guide processing by focussing its resources on the more interesting aspects of the story. Since the heuristics are geared towards learning, this ensures that AQUA spends its time on those aspects of the story that are most likely to result in something useful being learned. Without its interestingness heuristics, AQUA would still learn the same things, but it would spend a lot more time drawing inferences that ultimately turn out to be irrelevant.

References

[Carbonell, 1979] J. G. Carbonell. *Subjective Understanding: Computer Models of Belief Systems.* Ph.D. thesis, Yale Universtiy, New Haven, CT, January 1979. Research Report #150.

[Hayes-Roth and Lesser, 1976] F. Hayes-Roth and V. Lesser. Focus of attention in a distributed logic speech understanding system. In *Proceedings of the IEEE International Conference on ASSP*, Philadephia, PA, 1976.

[Hidi and Baird, 1986] S. Hidi and W. Baird. Interestingness — A Neglected Variable in Discourse Processing. *Cognitive Science*, 10:179–194, 1986.

[Lebowitz, 1980] M. Lebowitz. *Generalization and Memory in an Integrated Understanding System.* Ph.D. thesis, Yale University, Department of Computer Science, New Haven, CT, October 1980. Research Report #186.

[Ram, 1987] A. Ram. AQUA: Asking Questions and Understanding Answers. In *Proceedings of the Sixth Annual National Conference on Artificial Intelligence*, pages 312–316, Seattle, WA, July 1987. American Association for Artificial Intelligence, Morgan Kaufman Publishers, Inc.

[Ram, 1989] A. Ram. *Question-driven understanding: An integrated theory of story understanding, memory and learning.* Ph.D. thesis, Yale University, New Haven, CT, May 1989. Research Report #710.

[Schank and Ram, 1988] R. C. Schank and A. Ram. Question-driven Parsing: A New Approach to Natural Language Understanding. *Journal of Japanese Society for Artificial Intelligence*, 3(3):260–270, May 1988.

[Schank, 1979] R. C. Schank. Interestingness: Controlling Inferences. *Artificial Intelligence*, 12:273–297, 1979.

[Sperber and Wilson, 1986] D. Sperber and D. Wilson. *Relevance: Communication and Cognition.* Language and Thought Series. Harvard University Press, Cambridge, MA, 1986.