Artificial Intelligence Review (1991) **5**, 239–254

Understanding and summarization

Richard Alterman
Computer Science Department, Brandeis University, USA

Abstract. This article is an overview of the literature on narrative summarization. The capacity to summarize is a fundamental property of intelligence and has significance for several areas of artificial intelligence research and development. The first part of the paper includes a description of four critical features of a summary. The bulk of this review is concerned with sorting available summarization frameworks and techniques. A latter section of the paper describes the significance of summarization technology to three current topics in artificial intelligence: explanation-based learning, case-based reasoning, and plan evaluation.

1 Introduction

This paper is a review of the literature on summarization. The basic idea of summarization is to take a body of information and reduce its size and content to its important points. The capacity to summarize is a fundamental property of intelligence. In our daily lives there is an overwhelming amount of information to process and much of it is neither relevant nor of interest. Summarization processes allow an intelligent agent to focus on the most significant aspects of a given understanding.

A pragmatic reason for studying summarization is that it provides a useful way to report back on a body of knowledge. Take a simple example: suppose a user sits down to an information retrieval system and requests all information relevant to Japanese embargos of imports. In a given knowledge-base there might be an overwhelming number of episodes to describe. What the user does not want is a detailed description of each episode. Rather, s/he would prefer a summary of the more important episodes.

Summarization also serves an important function in research on cognition. One of the central issues of cognitive science is characterization of the 'understanding' process. Summarization provides a test of a given model of understanding. Suppose a researcher claims that the thematic level of understanding plays a significant role in understanding. Summarization provides a test and a methodology for exploring this sort of claim. If thematic understandings are important, then they will be reflected in the sorts of summaries that human subjects produce for a given text and thus should also be reflected in summaries produced by computer models.

Summarization is not a single phenomenon. There are many different kinds of summaries. To name just a few, there are abstracts, epitomes, overviews, abridgements, digests, and recapitulations. Each style of summarization requires a slightly different viewpoint on extracting the essential content of a given text or understanding.

Different approaches to summarization emphasize alternate effects and functions of the process. Summarization is sometimes treated as a problem of memory (e.g. Kintsch & van Dijk, 1978) i.e., what does the subject remember of the text after various periods of time. The point here is that as, for example, a story is read it is not perfectly stored in memory, but only its most significant parts are retrievable; by what process does this occur? Other models of summarization are biased towards one or other implicit structure of the text. Structures that have been previously computationally tested for their contribution to the summarization process include: story schema (Rumelhart, 1975), schema narrative trees (Simmons & Correira, 1980; Correira, 1980), sketchy scripts (DeJong, 1979), plot units (Lehnert & Loiselle, 1989; Lehnert, 1981), story points (Wilensky, 1980), and generic knowledge structures (Graesser & Clark, 1985). Each of these structures presents a different point of view on the underlying understanding and summarization processes.

A first pass approximation of some important properties of a summary, includes the following:

- does the summary reduce the workload for the interpreter/understander over the text?
- does the summary maintain coherence?
- does the summary maintain coverage?
- does the summary include the important events of the story?

The issue of *workload* suggests that it should take less work to construct an interpretation of a summary than the original text. (A summary is not only shorter but it is also, in some manner, 'simpler'.) The question of *coherence* suggests it is not good enough to just reduce the quantity of text — the summary must hold together and make sense. The third question (*coverage*) indicates that a good summary must cover, at least implicitly, many of the events of the original text. The last issue (*importance*) indicates that the summary should include the important parts of the text, and, where they are not necessary for reasons of coherence, exclude the unimportant parts.

The remainder of the article begins by expatiating on the axis role of representation in computational models of understanding and summarization. The bulk of the paper will be concerned with sorting the available summarization frameworks and techniques. Although traditionally summarization has been studied in the context of narratives, a latter section of this paper will describe why summariztion research is significant for three other areas of artificial intelligence: explanation based learning, plan evaluation, and case-based reasoning.

2 Representation, the artefact of 'understanding'

What does it mean to say that a machine has modelled 'understanding'? The traditional answer in artificial intelligence (AI) models of text comprehension is to say that a program that 'understands' can produce a representation. This representation includes not only the initial elements of the text but also includes a representation of the connections amongst these elements. Given the text (adapted from *The Peasant and the Waterman*, Protter, 1961):

> A peasant was chopping wood. He dropped his axed. It fell with a splash into the lake.

The task of the program is to build a representation that captures the connectiveness of the elements of the text. For example, the representation/understanding might include, in some manner, the connection between the peasant and the chopping (e.g. the peasant is the 'agent' of the chopping) or the connection between the 'chopping' and the 'dropping' (e.g. because the peasant was 'chopping' he was 'holding an axe', because he 'held an axe' he was able to 'drop it'). The representation/understanding might include only the 'notions' explicitly invoked by the text or it could include also those notions invited by the text.

Thus it is the task of the program/understander to take the elements of the text (phrases) and decide on a set of relationships between those elements.

Figure 1 shows the relationship between the program that constructs the representation/understanding and the 'summarizer'. The input to the 'understander' is typically either the text in natural langue form, a syntactic parse of each of the sentences in the narrative, or each of the clause of the input in case notation form.

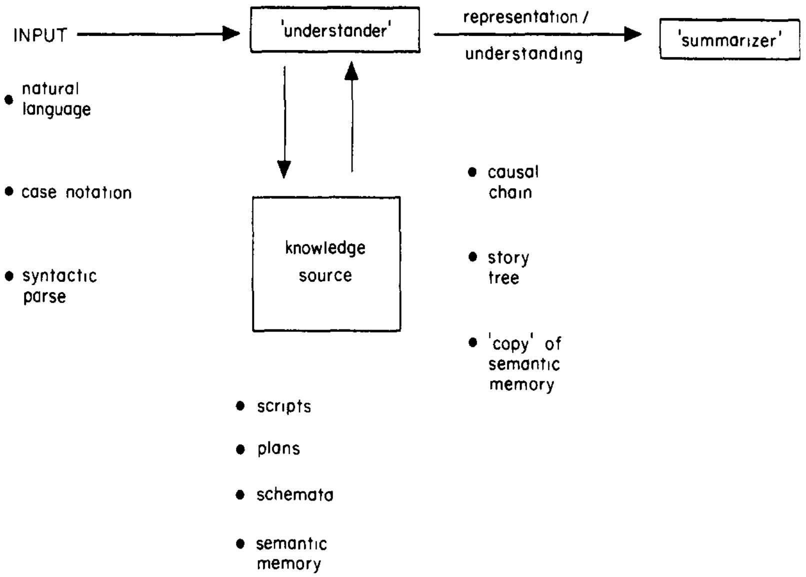


Fig. 1. Understanding/representation and summarization.

Given the input, the 'understander' uses some available knowledge source to build a representation of the coherence of the text. Examples of knowledge sources are: scripts and plans (Cullingford, 1978; Schank & Abelson, 1978; Wilensky, 1978), schemata (Grasser & Clark, 1985), and semantic memory (Alterman, 1985; van Dijk, 1977).

The actual form of the representation/understanding that is produced varies a great deal. In some cases it is a causal chain (e.g. Schank, 1977), in other cases a story tree (e.g. Rumelhart, 1975), in other cases a network representing the connectivity of co-referring expressions in the input text (Lockman & Klappholz, 1980), or a collection of copies of portions of semantic memory (e.g. Alterman, 1989).

There exists a tension between the ease of computing a given representation and its utility for such tasks as summarization and question answering; representations/understandings that are rich from the perspective of summarization are potentially difficult to produce. One of the reasons that summarization is important is that it can be used to evaluate the advantages, disadvantages, and differences amongst these various representation schemes. The represention/understanding both shapes and constrains the quality of a given summary. The summary shows the pertainence and efficacy of a given style of representation/understanding. In many cases summarization can be used to tune the knowledge source from which the understanding/representation was derived.

Given this framework for talking about understanding, each of the properties described in the introduction — coverage, coherence, importance and workload — can be grounded computationally. *Coverage* and *importance* are tied to features of the representation/understanding. The issue now becomes: to what extent does the internal structure of the representation/understanding present methods and explanations that account for coverage and importance? So, for example, if the representation/understanding forms a tree each level of the tree, arguably covers the entire understanding/representation at a different level of detail (e.g. Simmons & Correira, 1980; Simmons & Chester, 1982). Or, if the representation/understanding is a graph, those nodes either centrally located or maximally connected are candidates for important nodes.

The same mechanism that built the original representation/understanding can be used to build a representation/understanding of the summary. The summary lacks *coherence* if the 'understanding' mechanism fails to build an adequate representation/understanding of the summary. Finally, the reduction of *workload* can be evaluated by attaching a quantitative measure to the 'understanding' mechanism and comparing the effort it took to construct an interpretation of the original text, to the effort for constructing an interpretation of the summary.

3 In-the-small techniques

In-the-small approaches to summarization apply local techniques. They do not require an understanding of the text as a whole, but instead preserve the text's message and proportions by means of systematic abbreviations. They consider each piece of text in relative isolation and attempt to summarize.

Consider the following piece of text from *William Tell, the Archer* (Protter, 1961).

> Just then the clatter of horses' hooves was heard. And Gessler, the governor general, galloped into the square. His military retinue followed him. He reined his horse to a stop before the pole.

An in-the-large summary would require that the importance of this passage to the message of the text as a whole should be determined. An in-the-small summary applies local techniques to reducing the volume of this piece of text while maintaining its central content. A reasonable summarization-in-the-small of this text would be:

> The governor general rode into the square.

The summary includes the central event concept of this piece of text while deleting the fact that the clattering of hooves could be heard and the details of the riding. Notice that the event of 'riding' that is mentioned in the summary is not explicitly mentioned in the original text.

The early macro-structure rules of van Dijk (1977) are an example of in-the-small summarization techniques. An example of a macro-structure rule is *generalization*; concepts are generalized by abstracting them. So 'John is hitch-hiking' would be generalized to 'John is travelling'. Another rule is *deletion*, which abstracts accidental properties. So 'Mary was playing with the red ball' could be progressively summarized as 'Mary is playing with the ball' and then just 'Mary is playing'. A third example of macro-structure rule is *construction*, which replaces the components of a concept by the concept. So 'John laid the foundation. He built the walls . . .' would be summarized 'John built a house'.

Techniques that are tied to a particular knowledge structure can also be seen as in-the-small techniques. For example, FRUMP (DeJong, 1979) produces representations of text by applying in a top-down fashion sketchy scripts (e.g. accidents and terrorist acts). It extracts from wire-service newspaper stories just enough facts to fill in the arguments of a sketchy script. Because the stories that FRUMP works with are so stereotyped, it could summarize text by using a set of fill-in-the-blank type summarization statements attached to each sketchy script.

Here is an example taken from DeJong (p. 27: 1979):

> The Chilean government has seized operational and financial control of the U.S. interest in the El Teniente Mining Company, one of the three big copper enterprises here. When the Kennecott Copper Company, the owners, sold a 51 per cent interest in the company to the Chilean state Copper Corporation in 1967 it retained a contract to manage the mine. Robert Haldeman, Executive Vice President of EL Teniente, said the contract had been 'impaired' by the latest government action. After a meeting with company officials at the mine site near here, however, he said that he had instructed them to co-operate with eight administrators that the Chilean government had appointed to control all aspects of the company's operations.

FRUMP selected the 'NATIONALIZE' sketchy script, and it used it to produce the following summary: 'Chile has Nationalized an American mine.'

4 Three approaches to importance

4.1 The causal chain and importance

In 1975 it was suggested by Schank that for narrative texts the form that the representation takes is a causal chain (Schank, 1975). There have been two broad versions of the causal chain measure of importance: those events of greatest interest lie on the causal chain which corresponds to the major narrative thread (e.g. Schank, 1975; *critical path*; Black & Bower, 1980), and the other relates it to a count of the number of connections between a given node in the causal graph and other nodes in the causal graph (e.g. Graesser, 1981; Trabasso & Sperry, 1985; Trabasso & van den Broek, 1985).

Figure 2 shows a portion of text taken from 'The Father, His Son and Their Donkey (Trabasso & Sperry, p. 599, 1985). Figure 3 shows the underlying causal structure that they derived for this portion of the story. Trabasso & Sperry (p. 598) based this causal analysis on the Mackie (1980) test of causal necessity:

> Necessity is tested by the use of a counterfactual argument of the form: If not A then not B. That is, an event A is said to be necessary to event B if it is the case that had A not occurred then, in the circumstances, B would not have occured.

20. Do you see that idle boy riding the donkey.
21. while his father has to walk?
22. You should get down
23. and let your father ride!'
24. Upon this, the son got down from the donkey
25. and the father took his place.
26. They had not gone far
27. when they happened upon a group of women and children.
28. "Why, you lazy old fellow,
29. you should be ashamed."
30. cried several women at once.
31. "How can you ride upon the beast,
32. when that poor little boy can hardly keep up with you?"
33. So the good-natured father hoisted his son up behind him.

Fig. 2. A portion of 'The Father, His Son, and Their Donkey'.

For example, 22 is linked to 23 because, in the circumstances, if the boy had not dismounted from the donkey the father would not have mounted the donkey.

If we interpret critical path as the shortest path, according to the critical path notion of importance the most important events in the passage beginning at 21 and ending at 34 are: 20, 22, 24, 32, 34. Given the notion that the maximally connected nodes in a causal representation are the important ones, nodes 22, 24, 25, and 32 are the most important (each is connected to four other nodes in the graph).

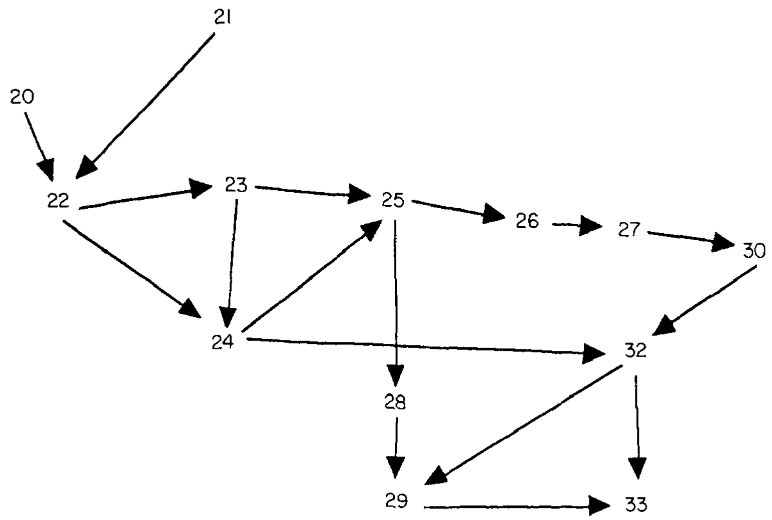


Fig. 3. The causal structure.

4.2 Thematic aspects of representation/understanding

Psychological evidence (Reiser, Black & Lehnert, 1985) suggests that readers do some inferencing on the thematic structure of the text. Thematic structure corresponds to a highly abstract context independent representation of events (e.g. competition). Experiments by Reiser et *al.* were based on the plot unit analysis developed by Lehnert (1981). The basic experiment performed by Reiser *et al.* was to ask a group of subjects to read a number of short texts and then sort them according to similarity. The findings of this research was that the sortings of subjects correlated to the sorting predicted by an analysis of thematic structure that was determined via a plot unit analysis.

There have been several computational efforts related to the question of thematic structure, including: *plot units* (Lehnert, 1981), *points* (Wilensky, 1980), *TAU's* (Dyer, 1983), and *TOP's* (Schank, 1982). Below are described two theme-based approaches to the study of importance.

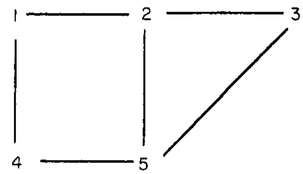
Lehnert (1981; Lehnert & Loiselle, 1989) developed a scheme for summarizing text based on plat units. Plot units represent affect-state patterns. Lehnert identifies a number of primitive plot units (e.g. motivation, success, perseverance) which can be combined into more complex plot units (e.g. fortuitous problem resolution, fleeting success, giving up). Narrative text is represented by interconnected plot units and summaries are based on the identification of pivotal plot units, i.e. the plot units which are maximally connected.

Here is an example, 'The Czar's Three Daughters', taken from Lehnert & Loiselle (1989, p. 147; also Graesser & Clark, 1985).

> Once there was a Czar who had three lovely daughters. One day the three daughters went walking in the woods. They were enjoying themselves so much that they forgot the time and stayed too long. A dragon kidnapped the

three daughters. As they were being dragged off they cried for help. Three heroes came and fought the dragon and rescued the maidens. Then the heroes returned the daughters to their palace. When the Czar heard of the rescue, he rewarded the heroes.

The resulting top level plot-unit graph is shown in Figure 4. The maximally connected nodes are: (1) competition between the heroes and dragon, and (2) the honored request between the daughters and the heroes.



1. Competition (Daughters and Dragon)
2. Competition (Heroes and Dragon)
3. Honored-Request (Daughters and Heroes)
4. Intentional Problem Resolution (Daughters)
5. Reward (Czar and Heroes)

Fig. 4. Thematic structure.

Wilensky introduced a model of thematic structure based on a model of interacting goals (Wilkensky, 1983). *Story points* (Wilensky, 1980, 1982) roughly correspond to the essential tension points of a story, i.e. what the story is about. The idea was that points represent what is interesting in a story (and therefore likely to be included in a summary). Each story point corresponds to one of a set of *point prototypes*. Wilensky argued that situations where goal interactions occur are potentially dramatic and consequently likely candidates as story points. An example point prototype is: *goal subsumption termination prototype*. Wilensky describes goal subsumption as referring to a situation in which a character's plan is to achieve a state that will make it easier for a character to fulfill a recurring goal. A dramatic situation occurs when a subsumption state is terminated (i.e. a goal subsumption point prototype). An example of goal subsumption state termination occurs in 'The Xenon Story' (Wilensky, 1983): When John loses his job he can no longer afford many of the things to which he had become accustomed.

4.3 *Knowledge-triggered and value-triggered importance*

Hidi & Baird (1986) and more recently Ram (1990; 1989) differentiate importance and interest techniques that are derived from a particular knowledge source from those techniques that characterize the unique perspective of the individual reading the text. (Hidi & Baird use the terminology *knowledge-triggered* versus *value-triggered*.) The value-triggered importance heuristics introduce the subjective bias of the reader.

The dissertation work of Ram (1989) computationally models value-triggered importance heuristics. Examples of the subjectively important portions of a text are those that are relevant to the reader's goal's or configuration of goals (or a are similar to them), situations that are anomalous from the perspective of the reader, contradict the reader's hypothesies, or present gaps in the explanation scaffolding that the reader brings to bear on the text. Or conversely, those parts of the text that are uninteresting to the individual reader are those that are not relevant to his/her goals or configurations of goals, are mundane, and have situations that are easily explained away.

Another example of value-triggered importance and interpretation can be found in the work of Gamson (1988), which explores how the cultural context provides a background against which understanding occurs. Gamson argued that cultural provides *interpretive packages* that frame an interpretation and function causally. He explains the following example. In the mid-1960s there was a partial meltdown in Detroit, yet it effected nuclear energy policy very little. Gamson asks the question: why is it that a decade later the same sort of incidence at Three Mile Island has such a major impact on nuclear energy policy? His argument is that it had to do with the interpretive packages. In the mid-1960s the sort of interpretive package available to describe the partial meltdown in Detroit was *faith in progress*, i.e. one step back for every two steps forward. Given such an interpretive package, the important feature of the incident in Detroit is that the partial meltdown is a *step back*, to be followed, in the future, by a *step forward*. By the mid-1970s, when the incident at Three Mile Island occurs, that are a lots of other competing interpretive packages available to interpret the same sort of event, including: *small is beautiful* and *public accountibility*. Given these sorts of interpretive packages different feature of the same kind of event become important, e.g. nuclear power plants are large-scale operations, or failing nuclear plants are excessively dangerous to the community.

5 The conceptual roots

Alterman & Bookman (in press) explore the encoding relationship between semantic memory and text (*event concept coherence*: Alterman, 1985; 1989). Semantic memory provides the basic vocabulary of concepts and 'practices', and their associated structure, for a given domain. The encoding relationship between semantic memory and text posits that the representation/understanding of the text is partially encoded by the vocabulary and structure provided by semantic memory. An important feature of this that semantic memory and the representation/understanding share the same structure; the structure of the representation/understanding is a 'copy' of some piece of semantic memory. This 'copy-based' notion of the representation/understanding is characteristic of semantic network-derived text representation schemes (e.g. Norvig, 1989; Martin & Riesbeck, 1986; Charniak, 1983, 1986; Alterman, 1985).

Consider the following description of events:

> The peasant (1) took the chest of gold to the czar. He (2) walked to the czar's majestic castle. In front of the czar's chambers, he was (3) halted by a haughty general.

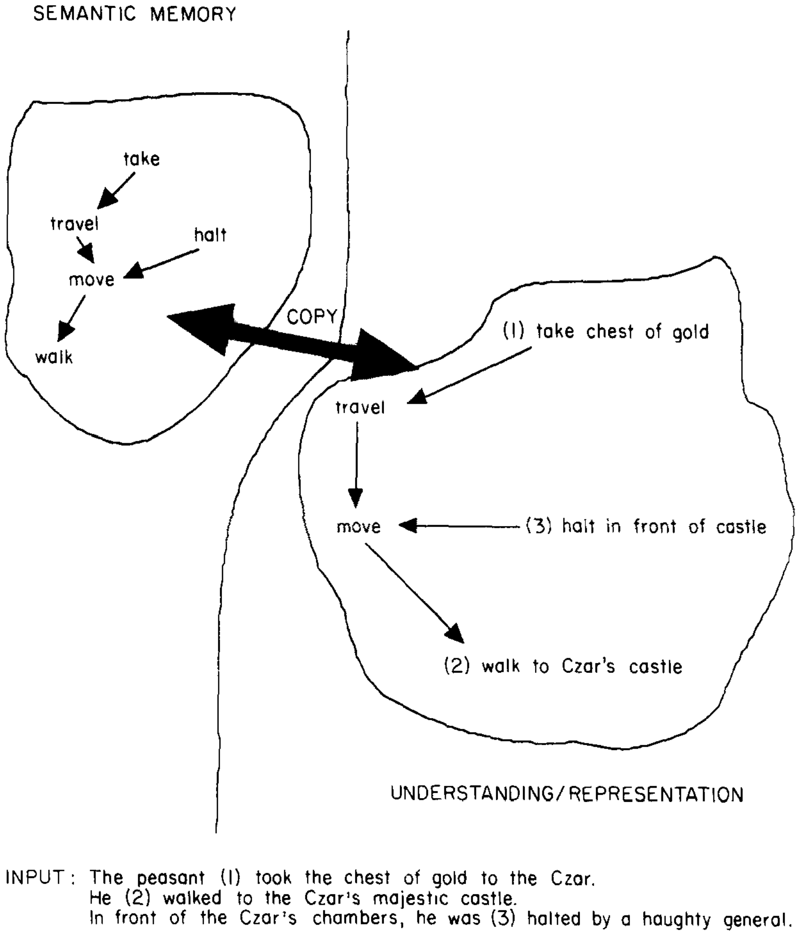


Fig. 5. The encoding relationship between semantic memory and the representation/understanding.

Figure 5 (taken from Alterman & Bookman, in press) depicts the representation/understanding of the text and its 'copy' relationship to a piece of semantic memory.

A program called SSS was developed to explore the utility of the encoding relationship between semantic memory and text. An important idea developed in the work on SSS was the notion of 'conceptual roots'. Roughly the conceptual roots correspond to the basic notions of the narrative text — the framework in the terms of which the narrative was developed. SSS determines the conceptual roots from a directed acyclic graph structure that represents the coherence of the event concepts in the text as derived from the encoding provided by semantic memory. An interesting property of the conceptual roots are that they are the minimal set that covers the semantic memory-based graph encoding of the case. For the piece of text shown in Figure 5 the conceptual roots are 'taking a chest of gold to the czar' and 'being halted in front of the czar's chambers'.

With the addition of the techniques that determine the conceptual roots, SSS is able to explain succinctly the connection between any two concept coherent events in the narrative. Also implemented in SSS is a measure of importance that quantifies the conceptual emphasis of the case. Given this measure of importance, it is easy to show that each of the important nodes is either a conceptual root or covered by one of the conceptual roots. Lastly, SSS combines evidence from semantic memory-based coherence graph of the case, the conceptual root analysis and the importance measure, to generate a description of the basic event content of the narrative (i.e. a basic summary).

The basic summary technique developed in SSS is as follows:

1. apply the conceptual root extraction technique in order to generate a list of event concepts in the story;
2. for each event in the list determine its relative importance using the coverage-based importance techniques described earlier;
3. remove from the list of events those events that have less than average importance.[1]

Because of the nature of the conceptual root list, this summary guarantees a fair degree of *coverage* — the events deleted are those that have few reachable nodes (i.e. little coverage). This technique also guarantees *importance*. Because importance (a count of reachable nodes) correlates with the coverage, and the conceptual roots maximize coverage, the conceptual roots will include, either implicitly or explicitly, the important nodes of the text. In order to demonstrate *coherence* and test the amount of *simplification* Alterman & Bookman ran the summaries produced by SSS through NEXUS (Alterman, 1985), the program which produced the original interpretation. Attached to NEXUS was a workload measure (Alterman & Bookman, 1990) that was used determine the amount of simplification in the interpretive process after summarization.

6 Is the representation/understanding a tree?

There exists a fair-sized body of work on narratives which attempts to represent part of the content of a narrative by looking at the underlieing syntactic structure of the narrative. (One would expect the structure of a mystery to be different than that of a regency romance.) Typically these types of narrative representations include story categories like 'episode' or 'setting' or 'episode resolution'. Overall these kinds of representations form a tree. There has been some work that has attempted to establish the primacy of certain categories as tending to be more important (e.g. Stein & Glenn, 1981; Mandler & Johnson, 1977).

There has also been some work on exploiting these tree-like representations for summarization purposes. A interesting property of such representation is that each level of the tree arguably provides *coverage* but at a different level of detail. Both Rumelhart (1975) and Simmons & Correira (1980; Correira, 1980) summarized text by level of the tree. Rumelhart worked with two trees, one contains the syntactic structure of the story the other its semantic structure. His system summarized the text by simultaneously descending both trees, deleting subtrees according to a set of semantic summarization rules. Simmons & Correira worked with a single tree which represented a combination of the syntactic and semantic structure of the story; the work of Kintsch & van Dijk (1978) describes a similar approach to text representation, where syntactic *superstructures* organize the semantic *macro-structures* of the text. For Simmons & Correira, any level of their tree represented a summarization of the story. No extra rewrite rules were required.

The following is a section of the text of the Black and Yellow V-2 Rocket (p. 154 of Simmons, 1984):

> . . . With a great roar and burst of flame the giant rocket rose slowly and then faster and faster. Behind it trailed sixty feet of yellow flame. Soon the flame looked like a yellow star. . . .

At one level of detail in the tree the ascent is described as follows (p. 192):

> The giant rocket rose with a great roar and burst of flame. It trailed sixty feet of yellow flame. Soon the flame looked like a yellow star.

leaving out the acceleration of the rocket. At still more abstract level of detail in the tree this segment of the story is summarized (p. 192):

> The giant rocket rose with a great roar and burst of flame.

7 A workload measure

Alterman & Bookman (1990) developed a measure of the work associated with a program reading a particular narrative. They used this measure to characterize the difference in work before and after summarization. The workload measure characterizes the average lifetime of an inference made by the 'understander' while building a coherence representation of the narrative. Texts with greater durations in the life of the average inference, reflect a higher complexity — greater amounts of work to read — because the reader must assume larger numbers of inferences over greater periods of time. Alterman & Bookman argue that their workload measure characterizes, along one dimension, how simple the text is in its construction.

Table 1 shows the relative thickness (as compared to the 'Margie Story', Rumelhart, 1975) associated with the analysis of each of several texts. It indicates that the 'The Clever Peasant and the Czar's General' is the most thick and the 'Margie Story' the least thick. The 'Restaurant Story', and 'V2-rocket Story', the

Table 1. The thickness complexity measure

Story	No. input events	Relative thickness
Margie	6	1.0
Restaurant	8	1.5
V2-rocket	9	1.9
Czar's daughters	13	2.3
Xenon	29	4.5
Clever peasant	93	42.1

'Czar's Daughters' Story' and the 'Xenon Story' follow in increasing order of complexity. Each of the first five texts were used as a vehicle for demonstrating one or another computational theory of summarization/understanding: 'The Margie Story' was used by Rumelhart (1975) for exploring story trees; 'The Restaurant Story' was used by Schank & Abelson (1977) for scripts; 'The Czar's Three Daughters' by Lehnert & Loiselle (1989) for plot units; 'The Black and Yellow V-2 Rocket' by Simmons & Chester (1982; c.f. de Beaugrande, 1980) for schema/narrative trees; 'The Xenon Story' by Wilensky (1980) for story points.

8 Other applications of summarization

Although much of the literature on summarization is concerned with the summarization of narratives, the scope of these results goes beyond narrations. Below I list some recent topics of interest in AI and cognitive science and describe how summarization comes into play.

Explanation-based learning. The idea behind explanation-based learning (EBL) is that learning can occur by a process of explanation. Given some *goal concept* (e.g. kidnapping), an EBL program *operationalizes* that concept by explaining a single example of that concept in the terms of some *domain theory* (e.g. Mitchell *et al.*, 1986; DeJong & Mooney, 1986). A critical part of the EBL enterprize is to determine the relevant features of the explanation of the example — and that is a problem of summarization. A second critical part of EBL learning is to determine the goal concept behind the example and importance suggest an approach to the detection of the 'goal concept'.

Case-based reasoning. The idea behind case-based reasoning is that reasoning can be supplemented, or even driven, by the usage of specific previous cases and/or episodes. An example of a domain where case based applies is the domain of legal reasoning, where the legal precedents act as the cases (Rissland & Ashley, 1986). Another domain where case-based reasoning has been explored is in the domain of foreign policy (Kolodner & Simpson, 1989); various international episodes play the role of cases. Summarization plays a role in case-based reasoning in a number of ways. For example, the whole question of importance, which is central to summarization, also has a significant role in case-based reasoning, as that the important features of a case are likely candidates for the 'indexing' features that are used to retrieve that case during the reminding phase of case-based reasoning. Also summarization techniques can play a key role in determining the central parts of a given episode, either for storing cases or in their explanation.

Plan evaluation. For large planning systems there is a problem of evaluating the utility, efficiency, and correctness of a given plan. One approach to evaluation is to run the plan in some kind of simulated environment. For larger plans, that might involve the co-ordination of several different large-scale plans, summarization offers an approach to evaluation. A summary of the plan allows a user to rapidly evaluate the plan at a high-level, matching a proposed plan against the user's goals or circumstances that are not readily simulatable.

9 Summary remarks

This article is an overview of the literature on summarization. Four key features of summarization are *coherence* (the summary must make sense), *coverage* (the summary must cover the bulk of the original understanding), *importance* (the summary must include the important features and exclude the unimportant ones), and *workload* (the summary is a simplification).

One approach to summarization is to nibble away at the original understanding/representation using a small arsenal of in-the-small techniques.

There have been several approaches to the determining the important features of the 'understanding'. Each approach reflects a different implicit structure in the representation/understanding. A causal chain-based representation/understanding uses notions of causal centrality and critical path to determine importance. A thematic structure analysis emphasizes the role of thematic features of the text in determining importance. Techniques that introduce the subjective bias of the reader or society also offer approaches to importance.

The notion of conceptual roots offers an approach to summarization that testably accounts for each of the four properties coherence, coverage, importance, and simplification. The conceptual roots are derived from a semantic memory encoding of the connectivity of events in the narrative. Because of the shared structure between the semantic memory and representation/understanding the conceptual roots are determinable.

Various researchers have proposed representation/understandings that include not only a coherence analysis of the text based on a semantic analysis but also further organizations based on the syntactic character of the text type. Generally speaking, these sorts of representation schemes result in a tree-like structure. One proposed advantage of this sort of analysis is that various syntactic categories of story structure correlated in various degrees with importance. A second potential advantage is that each level of the tree provides a summary of the story at a different level of abstraction, and, because of the tree-form, each level arguably provides coverage.

Although much of the work in summarization has been applied to narrative summarization, many of the ideas and techniques generated in this community have wide application to other areas of research and development in artificial intelligence. Three such areas were briefly discussed: explanation-based learning, case-based reasoning, and plan evaluation.

Notes

[1]Although we have not tested it here, an alternate scheme would continue to delete important events from this list until the summary loses coherence.

Acknowledgements

This work was supported in part by the Defense Advanced Research Projects Agency, administered by the US Airforce Office of Scientific Research under contract F49620-88-C-0058.

References

Alterman, R. (1985) A dictionary based on concept coherence. *Artificial Intelligence*, **25**, 153–86.

Alterman, R. (1989) Event concept coherence. In *Semantic Structures: Advances in Natural Language Processing*, (ed. D. Waltz) pages 57–87. Lawrence Erlbaum Associates, Hillsdale, NJ.

Alterman, R. & Bookman, L. (1990) Some computational experiments in summarization. Discourse *Processes*, **13**, 143–74.

Alterman, R. & Bookman, L. (1992) Reasoning about a semantic memory encoding of the connectivity of events. *Cognitive Science*, in press.

Alterman, R. & Zito-Wolf, R. (1990) Planning and understanding: Revisited. In: *Proceedings of 1990 AAAI Spring Symposium*.

Black, J. & Bower, G. (1980) Story understanding as problem solving. *Poetics*, **9**, 223–50.

Charniak, E. (1983) Passing markers: A theory of contextual influences in language comprehension. *Cognitive Science*, **7**: 171–90.

Charniak, E. (1986) A neat theory of marker passing. In: *Proceedings of the Fifth National Conference on Artificial Intelligence*, pages 584–8. Philadelphia, PA, Morgan Kaufmann.

Correira, A. (1980) Computing story trees. *American Journal of Computational Linguistics*, **6**, 135–49.

Cullingford, R. (1978) Script application: Computer understanding of newspaper stories. *Technical Report 116*, Yale University.

De Beaugrande, R. (1980) *Text, discourse, and process: Toward a multidisciplinary science of texts*. Advances in Discourse Processes (Vol. 4). Ablex Publishing, Norwood, NJ.

DeJong, G. & Mooney, R. (1986) Explanation-based learning: An alternative view. *Machine Learning*, **1**, 145–276.

DeJong, G. (1979) Prediction and substantiation: A new approach to natural language processing. *Cognitive Science*, **3**: 251–73.

Dyer, M. (1983) *In-depth understanding*. Cambridge MA, MIT Press.

Gamson, W. A. (1988) Political discourse and collective action. In: *From Structure to Action: social movement participation across cultures* (eds B. Klandermans, H. Kriesi & S. Tarrow). JAI Press.

Graesser, A. (1981) *Prose comprehension beyond the word*. Springer-Verlag.

Graesser, A. & Clerk, L. (1985) *Structure and procedures of implicit knowledge*. Ablex, Hillsdale, NJ.

Hidi, S. & Baird, W. (1986) Interestingness — a neglected variable in discourse processing. *Cognitive Science Journal*, **10**, 179–94.

Kintsch, W. & Van Dijk, T. (1978) Toward a model of text comprehension and production. *Psychological Review*, **85**, 363–4.

Kolodner, J. L. & Simpson, R. L. (1989) The MEDIATOR: Analysis of an early case-based problem solver. *Cognitive Science*, **13**, 507–49.

Lehnert, W. (1981) Plot units and narrative summarization. *Cognitive Science*, **5**, 293–331.

Lehnert, W. & Loiselle, C. (1989) An introduction to plot units. In: *Advances in Natural Language Processing* (ed. D. Waltz), p. 88–111. Lawrence Erlbaum Associates.

Lockman, A. & Klappholz, A. D. (1980) Toward a procedural model of contextual reference resolution. *Discourse Processes*, **3**, 25–71.

Mackie, J. L. (1980) *The cement of the universe: A study of causation*. Oxford University Press, Clarendon Imprint, London.

Mandler, J. & Johnson, N. (1977) Remembrance of things parsed: Story structure and recall. *Cognitive Psychology*, **9**, 111–251.

Martin, C. & Riesbeck, C. (1986) Uniform parsing and inferencing for learning. In: *Proceedings of AAAI-86*, pp. 257–61.

Mitchell, T., Keller, R. & Kedar-Cabelli, S. (1986) Explanation-based generalization: A unifying view. *Machine Learning*, **1**, 47–80.

Norvig, P. (1989) Marker passing as a weak method for text inferencing. *Cognitive Science*, **13**, 569–2620.

Ram, A. (1989) *Question driven understanding. An integrated theory of story understanding, memory and learning*. PhD thesis, Yale University.

Ram, A. (1990) Knowledge goals: A theory of interestingness. In: *Proceedings of the Twelfth Annual Conference of the Cognitive Science Society*, pp. 206–14, New Haven, CN, LEA Associates.

Rissland, E. & Ashley, K. (1986) Hypotheticals as heuristic device. In: *Proceedings of the Fifth National Conference on Artificial Intelligence*.

Rumelhart, D. (1975) Notes on a schema for stories. In: *Representation and understanding*, (eds D. Bobrow & A. Collins) pp. 211–36. Academic Press, New York, NY.

Schank, R. & Abelson, R. (1977) *Scripts, Plans, Goals, and Understanding*. Lawrence Erlbaum Associates, Hillsdale, NJ.

Schank, R. (1975) The structure of episodes in memory. In: *Representation and understanding*, (eds D. Bobrow & A. Collins) pp. 237–72. Academic Press, New York, NY

Schank, R. (1982) *Dynamic memory* Cambridge University Press.

Simmons, R. & Correira, A. (1980) Rule forms for verse, sentences, and story trees. In: *Associative networks: The representation and use of knowledge in computers*, (ed. N. Findler) p. 363–91. Academic Press, New York, NY.

Simmons, R. (1980) Rule forms for verse, sentences, and story trees. In: *Associative networks: The representation and use of knowledge in computers* (ed. N. Findler) pp. 363–91. Academic Press, New York, NY.

Simmons, R & Chester, D. (1982) Relating sentences and semantic networks with procedural logic. *CACM*, **25**, 527–47.

Simmons, R. (1980) *Computations from the English*. Prentice Hall Inc., Englewood Cliffs, NJ.

Stein, N. & Glenn, C. (1981) *An analysis of story comprehension in elementary school children. Advances in Discourse Processes, Vol. 2. Ablex Publishing.*

Trabasso, T. & Sperry, L. (1985) Causal relatedness and importance of story events. Journal of Memory and Language, **24**, 595–611.

Trabasso, T. & Van den Broek, P. (1985) Causal thinking and representation of narrative events. *Journal of Memory and Language*, **24**, 612–30.

Van Dijk, T. (1977) Semantic macro-structures and knowledge frames in discourse. In: *Cognitive Processes in Comprehension* (eds M. A. Just & P. A. Carpenter). Lawrence Erlbaum Associates, Hillsdale, NJ.

Van Dijk, T. (1980) *Macrostructures: an interdisciplinary study of global structures in discourse, interaction, and cognition*. Lawrence Erlbaum, Hillsdale, NJ.

Van Dijk, T. (1976) Macro-structures and cognition. In: *Proceedings of twelfth annual Carnegie symposium on cognition*.

Van Dijk, T. & Kintsch, W. (1983) *Strategies of discourse comprehension* Academic Press, New York, NY.

Wilensky, R. (1978) *Understanding goal-based stories*. PhD thesis, Yale University.

Wilensky, R. (1980) What's the point? In: *Proceedings of third national conference of the Canadian society for the computational studies of intelligence*.

Wilensky, R. (1982) Points: a theory of the structure of stories in memory. In: *Strategies for natural language processing* (eds W. Lehnert & W. Ringle). Lawrence Erlbaum Associates, Hillsdale, NJ.

Wilensky, R. (1983) *Planning and Understanding*. Addison-Wesley.