NARRATIVE COMPLEXITY BASED ON SUMMARIZATION ALGORITHMS*

Wendy G. Lehnert

Department of Computer and Information Science
Universtiy of Massachusetts
Amherst, MA 01003

- Abstract -

Narrative structures can only be defined in terms of some internal memory representation, but narrative complexity is more properly characterized by information processing requirements. Story grammars, plan and goal hierarchies, and causal chain representations all provide a sense of structure which is largely removed from the processes that produce or access that memory representation. In this paper we introduce the notion of algorithmic equivalence as a means of generating more algorithmically-oriented taxonomies for memory representations. Using memory representations based on plot units, we define two narratives to be algorithmically equivalent if they can be effectively summarized by the same retrieval process. This perspective on representational strategies is an especially natural one from a processing point of view, since the computational complexity of a particular information processing task must be measured in terms of the algorithms involved .

1. The Plot Unit Approach to Summarization

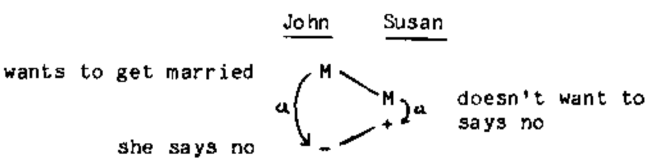
A representational strategy for narrative text has been developed to account for summarization behavior using relatively simplistic retrieval algorithms. When the memory representation for a narrative is encoded in terms of plot units [Lehnert 1980; 1981], it is possible to invoke retrieval algorithms that locate the central most important concepts of the narrative by examining structural features of cyclic graphs. Each node in the graph corresponds to a plot unit instantiation, and two nodes are connected by an arc when they share a common internal component.

A plot unit is a fixed configuration of smaller components called affect states. There are three affect state types designed to differentiate gross subjective states within a single character: positive states, negative states, and neutral mental states. These affect states emphasize emotional reactions to events and states rather than goal-oriented planning behavior, and each character in a narrative can be tracked in terms of an "affect state map" which chronologically records the subjective mental states for that character.

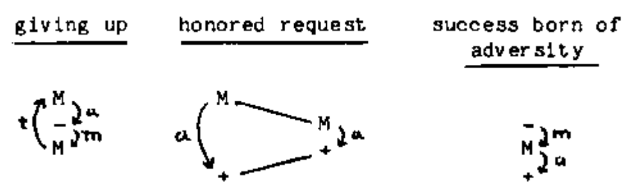
*This work is being supported by National Science Foundation grant //IST-8217502.

Once an affect state map has been produced which tracks the major characters of a narrative, we can look for instances of specific plot units within that representation. A "top level" plot unit instantiation is one that is not subsumed (fully contained) by any other plot unit instantiation. When all the top level units are recognized, we create a plot unit graph in which the nodes of the graph correspond to top level plot unit instantiations. Two nodes of the graph are then connected by an arc whenever they share at least one common affect state. This graph structure provides a level of memory representation that is especially well-suited for text summarization [Lehnert, Black, & Reiser 1981; Reiser, Lehnert & Black 1981; Gee & Grosjean 1982; Reiser, Black & Lehnert 1982; Lehnert, Alker, & Schneider 1983]. Nodes which are structurally central to this graph are expected to provide us with the conceptual content for a good summary.

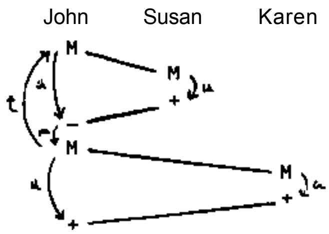
For example, suppose John asks Susan to marry him and she says no. This episode would be represented by the denied request unit:

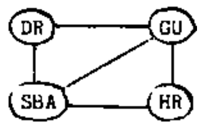


Now suppose John asks Karen to marry him and she says yes. This involves three more top level units:



If we check for overlapping plot units, we find that our plot unit graph of four units involves five arcs:





It is difficult to argue for the power of this representation when we are summarizing a two sentence story, but plot unit graphs can involve hundreds of units. For such complex narratives, we rely on structural features of the plot unit graph to tell us which nodes contain the concepts that are most central and critical to the story. For example, if a graph contains a unique node of maximal degree, we want to examine the contents of that node in order to produce a summary of the story. If a graph contains multiple nodes of maximal degree, we will have to resort to a different retrieval algorithm. We can therefore pursue a notion of narrative equivalence wherein two narratives are <u>algorithmically equivalent</u> (a-equivalent) if their plot unit graphs can be processed by the same summarization retrieval algorithm. It follows that narratives with isomorphic plot unit graphs are a-equivalent, although narrative plot unit graphs can be a-equivalent without being structurally isomorphic. This notion of algorithmic equivalence creates a weaker partition than the more traditional relation of isomorphic equivalence.

2. Narrative Equivalence Classes

Any partition of a-equivalent plot unit graphs determines a corresponding partition of narrative source texts. This in turn constitutes a taxonomy of narrative complexity. To identify such algorithmically-motivated equivalence classes, we must identify categories of graph structures that can be associated with specific retrieval algorithms. Any systematic development of such equivalence classes must be based on a large set of narrative representations that includes a wide variety of graph structures. This work is now in progress, but we have compiled enough narrative representations at the level of plot unit graphs to suggest some preliminary taxonomies. We are also developing a methodology for the identification of graph equivalencies which will allow us to proceed in a systematic manner.

The general strategy we are using to* discover plot unit graph equivalencies is a combination of bottom-up exploratory work and top-down hypothesis testing. In tue bottom—up pusss We collect and analyze random narratives, to create a library of plot unit graph representations. We will later draw from this library for substantiating evidence and counterexamples, but the first phase of our research is simple compilation. We are using narratives that come from AI work in natural language processing, cognitive psychology research on text comprehension, and published short stories by popular authors. The source texts range in complexity from a single paragraph to about 50 pages. Our most complex narrative thus far is Arnold Toynbee's synopsis of the story of Jesus.

For each narrative we must hand-code a chronologically-ordered representation of affect state maps. We cannot automate this process with any generality because the text processing techniques involved are highly domain-dependent and would require extensive knowledge encoding for each narrative attempted [see Dyer 1982 for a good discussion of what would be involved]. The graduate students who produce our hand-coded representations work together to assure uniformity in their encoding techniques, and we are developing encoding heuristics to aid others outside our immediate research group who would like to experiment with plot unit graphs.

Once a hand-coded affect state map has been produced, we process the representation with PUGG (the Plot Unit Graph Generator). Pugg is designed to accept any set of plot units as our set of universal unit structures, and any affect state maps that are consistent with the structural conventions for legal affect state configurations. PUGG returns an adjacency matrix for all top-level units along with other useful information about disjoint subgraphs and immediate unit families.

One of the important parameters in this work is the specification of a universa] set of top-level plot units. Graph structures will vary according to our selection of legal unit configurations, but we would ideally like to arrive at a taxonomy of graph structures that remains valid over a variety of universal sets. As a psychological theory of cognition and memory, we expect universal sets to vary across individuals. This variance could account for individual differences in summarization behavior as well as developmental differences between children and adults. It is therefore important to analyze each narrative with respect to more than one universal set, indexing each resulting graph with respect to its universal context.

To assure flexibility in this regard, we are analyzing each narrative with respect to three universal contexts. Context A is restricted to units involving no more than one or two characters. Context B is a subset of context A that contains only the simpler plot units of set A (most units in set B contain less than 10 affect states). Context C departs from sets A and B insofar as it contains units that involve more than two characters. Context C consequently contains plot unit configurations that are more complicated than those found in context A. We do not require all three encodings for any given story to be a-equivalent according to our partition of graphs. It seems quite plausible that some stories will be easier to summarize under one universal set than another (this should be especially true if one set is relatively impoverished).

In the course of compiling this library of representations, we are seeing some patterns emerge:

(1) Most graphs are fully connected.

(2) If a graph contains a unique node of maximal degree, it is probably small (containing < 15 units).

(3) Two distinct clusters that are strongly connected often describe the same events from different perspectives.

(4) If a graph can be partitioned into maximal clusters, boundary units between the clusters tend to be important.

With each new graph we generate, we must examine that graph to see where its critical nodes are located. Sometimes the critical nodes are structurally conspicuous. In these cases we can associate a plausible retrieval algorithm which appears to be appropriate for that graph. The same algorithm typically applies to a number of graphs, in which case we must identify necessary and sufficient conditions for the application of that algorithm. In other cases we may not be able to identify a simple algorithm, or a simple algorithm applies but does not produce a satisfactory summary. These apparent failures may force us to revise a previous summarization algorithm, or revise the necessary and sufficient conditions that signal the applicability of a particular algorithm.

The necessary and sufficient conditions that identify appropriate algorithms will define our graph equivalence classes. Hypothetical equivalence classes arise every time we propose a possible summarization algorithm, but we must be careful to maintain consistency throughout the system whenever a new class is proposed or an old class is altered. If we begin to amass a large number of equivalence classes, we will watch for hierarchical relationships among possible partitions. From a developmental perspective it seems quite likely that simple partitions might be refined into more complex partitions, and that such refinements would be associated with improvements in summarization behavior. With this in mind, we expect to find simplistic classifications that produce inferior summaries for some stories.

As our library grows, we will track error rates for each equivalence class as well as error rates for each set of universal plot units. In an ideal partition, the error rate across equivalence classes should be roughly uniform with respect to each universal context. On the other hand, the overall error rate across different universal contexts may vary considerably, in which case we will learn something about effective plot unit configurations. For example, if the error rate associated with context C is significantly higher than the rates for contexts A and B, we will have a strong argument against the inclusion of plot units involving more than two characters.

3. Graph Types and Summarization Algorithms

Thus far, we have analyzed about 20 narratives with respect to universal contexts A and B, and we have not yet formulated the units for context C. We need to look at more narratives before we propose a set of equivalence classes, but on the basis of this initial library we have identified three core equivalence classes which appear to produce reasonable summaries for a majority of the narratives. We will summarize these categories briefly.

A. Simple Graphs with Unique Pivots

One class of graphs exhibit unique nodes of maximal degree. While this class seems to be restricted to smaller graphs, we can reliably look to such pivotal nodes for the concepts most central to the narrative as a whole. This was the first algorithm we identified, and is therefore discussed in some earlier publications (e.g. see Lehnert 1981). While our initial work suggests that it is very difficult for a long narrative to fall into this category, we may see a higher frequency of stories in this class as our universal set of plot units expands. For example, it should be easier to create graphs of this type within Context C than Context A.

B. Complex Graphs with Multiple Pivots

This group of graphs is quite large and must be divided into smaller subsets for effective categorization. In some cases we see two nodes of maximal degree whose immediate families partition the entire graph into two subgraphs. The maximal nodes on the boundary of this partition then serve to give us a one-sentence summary for the whole story. In other cases, we see graphs where the maximal nodes themselves provide critical concepts for summarization. This is especially common when the two maximal nodes are adjacent to one another, in which case they are frequently of equal importance.

As we investigate this class further, we may find it necessary to assign relative weights to nodes which vary with the surrounding environment. For example, suppose a node of degree 10 has 9 neighbors that all have degree 6 (this sort of dense connectivity is rarely encountered). We may want to assign a lesser salience to such a node, favoring instead a node of degree 9 whose neighbors all have a degree of 2. A particularly elegant solution for these larger graphs would be to simply locate the subgraph composed of nodes with minimal eccentricity, where the eccentricity of a node is defined to be the largest distance from that node to any other node of the graph [Proskurowski 1980]. There are many such possibilities, and our initial explorations have not adequately differentiated their strengths and weaknesses.

C. Separable Graphs

This category contains large graphs (> 50 nodes) composed of subgraphs that can be separated by deleting a single node. While it may seem that the nodes central to each potential component of the graph might be good candidates for conceptual salience, we have found that the best results are obtained by looking at the "deletion" nodes which keeps the graph from separating. If there are multiple deletion nodes, we look for a path containing all the deletion nodes. The shortest path seems to be preferred, although maximal degree can enter in as a factor when more than one path is possible. These graphs tend to be associated with the longest and most complex narratives (see [Lehnert, Alker, and Schneider 1983] for a detailed discussion of one such graph).

*****

As we consider these three preliminary a-equivalence classes, we see that processing complexity is determined to some extent by the size of the memory representation (or plot unit graph). This is hardly surprising, although the concept of a-equivalence suggests that there should be distinct plateaus of summarization competence, with severe drops in performance whenever a representation is incorrectly categorized. This provides us with a set of performance predictions which will be distinct from any models that predict strictly linear complexities based on the length of the input text or the size of an internal memory representation alone.

4. Conclusions

Our work to identify a-equivalent narratives continues on many fronts. We are expanding our library of affect state maps and the corresponding library of plot unit graphs. At this stage it seems appropriate to concentrate primarily on the taxonomy of necessary summarization algorithms, and secondarily on issues concerning universal sets of plot units. We hope to concentrate more fully on the question of universal sets after we have a firmer footing on the issue of algorithmic equivalence. It is too early to say anything about the status of algorithmic equivalence as a concept of psychological import, but a plot unit approach to human text comprehension may dovetail very naturally with a developmental study of human summarization behavior. We might, for example, expect to see competent summarization behavior in some classes before others, in which case the notion of a-equivalence would provide a simple measure of narrative complexity in terms of human information processing capabilities.

On a more general level, we note that plot units are applicable to narrative texts with human (or anthropomorphized characters) for whom an affect state analysis is possible. It is not obvious that an analogous system of memory representation can be applied to expository or purely instructional texts. At the very least, our approach suggests that the cognitive processes underlying summarization skills for narratives is very different from the processes of summarization required for other types of text. Narratives and expository texts may therefore reside within two basically disjoint a-equivalence classes, in which case competence in one area may not correlate very strongly with competence in the other. Such a situation would have immediate consequences for educators as well as other computational models of text comprehension.

REFERENCES

Dyer, M. (1982). "In-depth understanding: a computer model of integrated processing for narrative comprehension." TR//219. Dept. of Computer Science, Yale University.

Gee, J. P., and Grosjean, F. (1982). "Empirical Evidence for Narrative Structure." (unpublished paper) Program in Applied Psycholinguisties, Boston University. Boston, Mass.

Lehnert, W.G. (1980). "Narrative Text Summarization." Proceedings of the First Annual National Conference on Artificial Intelligence. Palo Alto, California.

Lehnert, W.G. (1981). "Plot Units and Narrative Summarization." Cognitive Science. vol. 5, No. 4.

Lehnert, W., Alker, H., and Schneider, D. (1983). "Affective Plot Structure of Toynbee's Christus Patiens." Proceedings of the Sixth International Conference on Computers and the Humanities. Washington, D.C.

Lehnert, W., Black, J., and Reiser, B. (1981). "Summarizing Narratives." Proceedings *of* the Seventh Internation Joint Conference on Artificial Intelligence. Vancouver, British Columbia.

Proskurowski, A. (1980). "Centers of Maximal Outer planar Graphs." Journal of Graph Theory. Vol. 4; pp. 75-79.

Reiser, B., Black, J. and Lehnert, W. (1982). "Thematic Knowledge Structures in the Understanding and Generation of Narratives." Cognitive Science Technical Report //16. Cognitive Science Program, Yale University. New Haven, Ct.

Reiser, B., Lehnert, W., and Black, W. (1981). "Recognizing Thematic Units in Narratives." Proceedings of the Third Annual Cognitive Science Conference. Berkeley, California.