SUMMARIZING NARRATIVES

Wendv G. Lohnert, John B. Black, and Brian J. Reiser

Department ot Computer Science
Y.-tle University

ABSTRACT

Most research on narrative text summarisation has been conducted within the paradigm of experimental psychology. But recent language processing research in artificial intelligence suggests that the predominant theory of text summarisation requires further examination. Seemingly minor structural modifications of a story can reault in aignificant alterations of summary behavior. In this paper, highlights of summary data from 72 subjects are presented and analyzed in terms of two competing summarization models: (1) the story grammar model of psychology, and (2) the plot unit model developed in artificial intelligence. We will show how selected story grammar predictions compare to plot unit predictions for short term summarization and then identify two complicating factors that have a major impact on summarisation behavior.

1. Introduction

While experiments using people may seem less relevant to artificial intelligence than experiments using computers, the work described in this paper represents more than an intellectual commitment to psychological validity in AI (although that commitment is also present). The task of text summarisation is one of the most complicated challanges facing natural language processing and poasibly facing AI altogether. We need all the help we can get.

We have been working in recent years at Yale on a story understanding system (BORIS) (2] whose theoretical roots go back to the MARGIE system [10]. While an initial implementation toward summarisation has been completed in BORIS [7] , we have not yet been able to test our most important intuitions using BORIS. We probably won't be able to utilise BORIS for this purpose until it can process a large number of prototype stories. In the meantime, we have been conducting exploratory psychological experiments, to see how our ideas hold up against systems that can summarize a large number of stories. The results have been both surprising and encouraging.

This work was sponsored in part by the Advanced Research Projects Agency under contract N0014-75-C-1111 and in part by the National Science Foundation under contract IST7918463.

2. Plot Units: A Heterarchical Representation

Plot units (originally called "affect units") 13,4] lead to general graphs that are typically cyclic in their atructure rather than hierarchical. A plot unit ia a fixed configuration of mental states and events involving one or more characters in a atandard behavioral pattern or interaction. For example, there are plot units that represent honored requests, problem resolution, retaliation, broken promises, and obligations.

All plot units are constructed from smaller primitive components called "affect states." There are three kinds of affect states, designed to encode mentsl states, positive events, and negative events. The internal structure of a plot unit is governed by strict rules to control the various ways affect states can be linked to each other [3], but the most important property of a plot unit representation lies in the fact that plot units can overlap with each other. Two plot units overlap when they both contain a common affect state. For example, suppose John asks Mary to buy an artichoke, and she does so. The unit describing John's honored request will overlap with the unit describing Mary's success because they both share the same positive event of Mary buying an artichoke. This event is a critical component for both Mary's success and John's request:

John	Mary
wants Mary to buy artichoke	
	wants to buy artichoke
	buys artichoke
Mary buys artichoke	

Over lappings between top-level units provide us with the structure of a plot unit analysis. In the representational graph, nodes represent instantiated plot units, and arcs join two nodes whenever the corresponding plot units share a common affect state.

Each plot unit is associated with generational frames. Generational frames drive the natural language expression of instantiated plot units, and all units are associated with both strong (detailed) and weak (sketchy) generational frames. Stronger frames are used when there is strong connectivity and weaker frames are used when connectivity is weak. The process of summarizing a story is partly a function of structural connectivity, and partly a function of generational frame management.

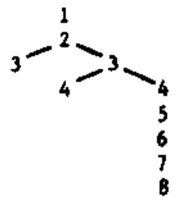
3. Three Storiea and Six Story Representations

To compare story grammars with plot units, we first designed three variations on a baseline story and collected 12 summaries on each variation for a total of 36 summaries (using 36 subjects). All experimental materials were presented as typed text. After reading the text, subjects were asked to produce a written summary in one or two sentences. The three stories (see Appendix A) are very cloae to each other in overall content. They each involve a deeply nested set of subgoals which the main character devises and achieves. The differences can be summarized as follows:

	Story 1	Story 2	Story 3
patio-building goal	Y	Y	N
hint dropping to Paul	Y	N	N
request to boss (denied)	Y	N	Y
request to CPA (initially denied)	N	Y	N

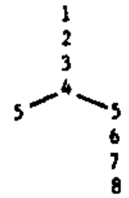
("Y" indicates the prescence of an item; "N" indicates its absence.) It is easy to generate simplified tree representations for the three stories to distinguish them structurally:

(1) BUILD PATIO
(2) GET PAUL OUT OF TOWN
(3) DROP HINTS TO PAUL (failed)
(3) GET BOSS TO SEND PAUL AWAY
(4) ASK BOSS (failed)
(4) GET ACCOUNTANT TO TRICK BOSS
(5) MAKE DEAL WITH ACCOUNTANT
(6) GIVE GRASS TO ACCOUNTANT
(7) GET GRASS FROM DEALER
(8) PAY UP TAB WITH DEALER



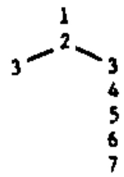
HIERARCHICAL TREE REPRESENTATION FOR STORY 1

(1) BUILD PATIO
(2) GET PAUL OUT OF TOWN
(3) GET BOSS TO SEND PAUL AWAY
(4) GET ACCOUNTANT TO TRICK BOSS
(5) ASK ACCOUNTANT (failed)
(5) MAKE DEAL WITH ACCOUNTANT
(6) GIVE GRASS TO ACCOUNTANT
(7) GET GRASS FROM DEALER
(8) PAY UP TAB WITH DEALER



HIERARCHICAL TREE REPRESENTATION FOR STORY 2

(1) GET PAUL OUT OF TOWN
(2) GET BOSS TO SEND PAUL AWAY
(3) ASK BOSS (failed)
(3) GET ACCOUNTANT TO TRICK BOSS
(4) MAKE DEAL WITH ACCOUNTANT
(5) GIVE GRASS TO ACCOUNTANT
(6) GET GRASS FROM DEALER
(7) PAY UP TAB WITH DEALER



HIERARCHICAL TREE REPRESENTATION FOR STORY 3

While not all propositions appear in all stories, the pairvise ordering of all propositions remains consistent throughout. It is never the case that proposition A is higher than proposition B in one tree, and lower in another. Story grammars predict inclusion on the basis of relative tree position. For example, the story grammar model predicts that proposition PI will appear more frequently than proposition P2 across all story versions:

PI: get boss to send Paul away
P2: Mike makes deal with accountant

Since we will be concentrating on such hierarchicial predictions, we should mention that the "critical path" model of Black and Bower predicts that failed goals and plans will violate the hierarchical inclusion principle in the sense that failure-associated nodes are unlikely to be mentioned no matter where they reside in the tree [1].

Plot unit predictions are based on the plot unit graphs which reflect a somewhat different set of structural distinctions. The units needed to graph our stories are:

PAF	perseverence after failure (signals the survival of a goal after failure)
DR	denied request (denial of a request)
SUB	nested subgoals (subgoal relationships where one goal is not achieved)
NS	nested successes (subgoal relationships where both goals are achieved)
RHC	request honored with condition (request honored with conditional bargaining)
NA	nested agencies (intermediary agents achieve nested subgoals)

While space forbids going into the details of the underlying affect state maps for each story, 11 plot units are instantiated for each story: To derive the plot unit graphs, arcs were inserted between two units whenever those units shared a common mental state or event.

SUB1	Mike gets Paul away to build patio
PAF1	Mike's persistence after 1st hint
PAF2	Mike's persistence after 2nd hint
NS1	Mike obtains boss's agency to get Paul away
NS2	Mike obtains CPA's agency to get boss's agency
NS3	Mike obtains grass so he can give grass to CPA
DR1	boss says no to request (in S1, S3)
DR2	CPA says no to request (in S2)
PAF3	Mike's persistence after denied request
RHC1	Mike's deal with CPA
RHC2	Mike's deal with Joe
NA1	Mike causes the accountant to cause the boss to cause Paul's trip

PLOT UNITS FOR STORIES S1, S2, and S3

Plot unit predictions are based on the connectivity or degree of each unit in the plot unit graph. Highly connected unite are sore likely to appear in summaries than weakly connected unite.

	S1	S2	S3
SUB1	4	2	(NA)
NS1	6	3	4
NS2	5	5	5
NA1	7	6	5
RHC1	5	5	5
NS3	2	2	2
RHC2	1	1	1

CONNECTIVITY CHART FOR PLOT UNITS

In keeping with the "critical path" hypothetic, we have omitted the failure-related unite (the DR's and the PAF't) eince we do not expect them to show up in the summaries. (Our data supported critical path prediction!).

To see exactly what theee connectivities predict, we suet consider the generational frames for each of the relevant units. Since plot units can be described with varying levels of detail, we will include a strong and a weak frame for each. Strong frames are more likely to appear when units are highly connected while weak frames are more likely to appear for weakly connected units.

SUB (subgoal)
Strong: X needed to do Y because X wanted to do Z
Weak: X wanted to do Z.

NS (nested successes)
Strong: X did Y in order to do Z
Weak: X did Z

RHC (request honored with condition)
Strong: X made a deal with Y in exchange for W
Weak: X made a deal with Y

NA (nested agencies)
Strong: X arranged to have Y get Z done
Weak: X arranged to get Z done

The strong generational frame for nested agency is meant to include the first agent in the chain, the agent cloaeat to X. In our stories, this agent is the accountant (rather than Paul'a boss who is further away from Mike in the eucceasful chain).

In determining whether a given unit is present in a summary, we naturally cannot expect to find verbatim generational framea. Some flexibility is needed to recognise verba like "bribe," "bargain," or "negotiate" as weak evidence of an RHC unit. In our stories, strong evidence of the RHC1 unit is represented by mentioning what Mike gave to the CPA (grata).

Some examples of inatantiated generational framea are given below:

SUB1(strong):

"Mike needed to get Paul out of town because he wanted to build Paul a patio for his birthday."

NSI(weak):

"Hike got Paul out of town."

RHCI(weak):

"Hike had successful negotiations with an accountant.

NAI(strong):

"Hike arranged with an accountant to have Paul sent out of town."

4. Selected Data Analysis

A comprehensive analysis of all the data would show that story grammars and plot units agree on many predictions. For example, both models correctly predicted that Hike's interaction with the drug dealer was least likely to show up in summaries. To tease apart the two models, we will examine summary data in four "target" areas where their predictions differ.

Target Area 1: "Hike gives graas to the accountant."

The story grammar predictions for this proposition are largely uniform across all three stories. This proposition appears at level 6 in SI and S2, and at level 5 in S3. This predicts that S3 summaries should mention Hike giving grass to the accountant more often than S1 and S2 summaries, with S1 and S2 mentioning it with equal freqency.

The plot unit analysis relies on our available generational frames. The only time the grass should appear within one of these frames is when the RHC1 unit is strongly present. From the three connectivity graphs, we see that RHC1 appears with varying strengths in the three stories:

	S1	S2	S3
RHC1	weak	moderate	strong

strong = Max
moderate = Max-1
weak < Max-1

These are relative strength rankings computed with respect to the maximal connectivity within each graph (Max(SI)- 7, Max(S2) • 6, Max(S3) - 5). The differences in relstive unit strength lead us to expect a difference in the frequencies of our target proposition for SI, S2, and S3. S3 summaries should include the target proposition more often than S2 summaries, and S2 ahould exhibit a higher frequency than S1. The data confirms the plot unit prediction:

	S1	S2	S3
give grass to CPA	.16	.50	.66

(Note: all percentile compariaions across stories in this paper pass the chi-equare test for significance at the .05 level unless otherwise specified).

Target Area 2: "The boat lends Paul away."

Again, the atory grammar predictions are fairly uniform. Mike's goal concerning Paul's boss appears at level 3 in SI and S2, and at level 2 in S3. This suggests that the proposition will be mentioned most often for S3, and slightly less often for S1 and S2, but with equal frequency for SI and S2.

The plot unit predictions are based on the connectivity of NS1 and NS2. A review of our generational frames and nested success unit instantiations shows that the target proposition only appears when NS1 is strongly present or NS2 is present (either weakly or strongly). Using the same computation of unit strength described above, we see that NS1 and NS2 appear with varying degrees of stength:

If the effects of NS1 and NS2 are taken into account with equal weight, we expect that summaries of S3 are most likely to mention the boss, with summaries of S1 and S2 somewhat less likely to mention it. Since the strength values in S1 and S2 cancel each other out, we would furthermore expect SI and S2 to exhibit equal fequencies. This analysis coincides perfectly with the story grammar analysis. But an alternative plot unit analysis points to a different prediction.

Notice that we need a strong generational frame for NS1, and NS1 appears with only moderate strength at best. If nested success units require strong connectivity for a strong generational frame, then all mention of the boss would be coming exclusively from NS2. If the effects of NS1 are discounted for not being strong enough, we would expect S2 summaries to mention the boss more often than S1 summaries since NS2 is stronger in S2 than in S1.

Both models predict that S3 summaries will describe the boss sending Paul away more often than the S1 and S2 summaries. But only a plot unit analysis can explain why S2 summaries might mention this more often than S1 summaries:

	S1	S2	S3
boss sends Paul away	.08	.33	.75

Target Area 3: "The boss sends Paul away." vs. "Mike makes a deal with an accountant."

In addition to comparing frequency distributions for single propositions across all three stories, we can also compare the relative frequency distributions for two propositione within each story. To see how two propositions should compare, we will look at their relative tree positions and graph connectivities:

	S1	S2	S3
boss sends Paul away (NS2) *	level 3 weak	level 3 moderate	level 2 strong
deal with CPA (RHC1)	level 5 weak	level 5 moderate	level 4 strong

Within each story the predictions are uniform. Story grammars predict that the deal with the accountant will be mentioned leas often than the boss sending Paul away because the deal occurs at a lower level in the tree. Plot units predict that these two propositions will be mentioned with equal frequency for each story version, since their connectivity rankings are identical in each story.

The data will show that the plot unit prediction is confirmed for S3, and neither model is confirmed for S1 and S2, although the plot unit predictions are closer to the data in these cases than the story grammar predictions:

	S1	S2	S3
boss sends Paul away	.08	.33	.75**
deal with CPA	.50	.58	.66**

In stories S1 and S2, The deal with the CPA appears significantly more often than the boss sending Paul away. It is also significant that the deal with the CPA appears with uniform frequency across all three stories*** since a plot unit analysis comparing the RHC1 unit across the three stories would not predict uniform frequency. A confounding factor which we will dub "generational subsumption" appears to be at work in S1 and S2. We will discuss this complication in section 5.

Target Area 4: "Mike wanted to build a patio."

This proposition did not appear in S3, but it was part of S1 and S2. A atory grammar analysis places this proposition at the head of the tree structure, which affords it the highest possible probability for inclusion. In the plot unit analysis, the patio goal appears only in the SUB unit which is ranked weakly for both stories. In S1, SUB is ranked with less strength than NS2 (which showed up in only 1 summary). In S2, SUB is ranked with the same strength as NS3 (which showed up in 3 summaries).

In this case, the story grammar predictions were clearly supported by the data. All 12 Si-subjects and all 12 S2-subjects mentioned the patio goal in their summaries. In fact, the first sentence in 21 of the 24 summaries sound exactly like clear-cut examples of the subgoal unit generational frame:

♦Since the data presented in Target Area #2 showed that NS1 had no influence over the frequency for the boas sending Paul away, we will discount NS1 here as well. The only unit that can carry the boas sending Paul away ia NS2.

**The difference between .75 and .66 is not significant.

*** The differences in .50,.58,.66 are not significant.

"Mike vented to make hit old friend Paul leave town for e few deye to thet he could build him a eurpriee present in hit

"Mike wanted to build hit friend Paul a patio for hie birthday, but he had to get Peul out of town for a couple of daya to do so."

"Mike wants to get his old friend Paul out of town for e few dayt to he can build a patio at Paul's house as a surprise•"

While the story greener model it one clear explanation for this overwhelming consentut, a self-containment factor could also be operating here. We will discuss this in the next section.

5. Generational Subsumption and Self-Containment

Both etory grammars and plot units attempt to predict summarization behavior in terms of internal memory representations.

The plot unit graph providee an important predictive element by revealing connectivity properties. But generetionel fremee must be teken into coneideration as well. For example, Mike's goal of getting Paul out of town was mentioned by ell 36 subjects, yet the NS1 unit which contains this goel in its top-level eucceee etructure is not strongly predicted in any of the plot unit graphs. However, the plot unit analysis is still consistent with the summary data aince the neated agency unit (NA) ia strongly predicted in all graphs, and NA cannot be expreesed without deecribing Mike's goal to get rid of Peul. In this case, we say that NA subsumes NS1 for the purposes of generation.

A similar form of generational subsumption across plot units occurs in the summaries for S1 and S2 when the deal with the accountant is mentioned. We saw in our discussion of Target Area #3 how the deal with the accountant was mentioned with uniform frequency across all three stories. This violated the different connectivity rankings that RMC1 assumed in S1, S2, end S3, along with the plot unit predictions for S1 and S2 deacribed in Target Area #3. But generational frames provide us with en explanation for thia. Conaider the generational frames for a request honored with a condition and a netted agency:

NA (strong): X arranged to have Y get Z done

RHC1 (weak): *X* made a deal with Y

Since the MA unit it strongly connected in S1 and S2, we know we have a commitment to the strong NA generational frame. We also have RHC1 weakly connected in S1 end moderately connected in S2. If we take the number of words involved at a rough indicator of complexity, it ia no harder to generete "NA(strong) + RHCl(weak)" than it is to generate "NA(strong)" alone:

MA(t):
Mike arranged to have an accountant
get Paul out of town.

NA(s) ♦ RHCl(v):
Mike made a deal with an accountant
to get Paul out of town.

While the effort involved in generating theee two concepts it not aignificantly different, there ii a significant difference in their content. "NA ♦ RHCl" is much more specific than "NA" by itself. This makes it optimally efficient to generate the weak frame for RHC1 whenever the strong frame for NA is being generated:

	S1	S2	S3
NA	.75	.66	.82
RHC1(w)	.58	.58	.66
RHC1(s)	.16	.50	.66

The strong frame for RHC1 is not pulled in by the NA unit, because it requires more effort; an additional phrase must be generated to express RHC1 strongly:

NA(s) ♦ RHCl(v):
Hike made a deal with an accountant
to get Paul out of town.

NA(s) + RHCKs):
Mike made a deal with an accountant
to get Paul out of town in exchange
for some grass.

Generational subsumption is not structural by nature. It is language dependent, and tightly bound up with the complexities of language generation.

Another confounding factor may be lurking behind the overwhelming presence of the subgoal unit, as discussed in Target Area #4. When people evaluate summaries, they are sensitive to how self-contained a summary sounds. It seems ressonable that this sensitivity to self-containment would enter into the summarization process as well.

While the plot unit analysis for S1 and S2 tells us that getting Paul out of town is the most importent top-level goal, this is not a very self-contained goal. Why would Mike want to get his friend out of town for a weekend? Are his motives good or bad? What's he up to? The goal begs for an explanation.

To aee what role self-containment might be playing in S1 and S2, we examined data from two additional stories that are structurally identical to S1 and S2 In these new stories, a farmer replaces Mike, and a donkey replaces Paul. Inatead of Mike wanting to build a patio, the farmer wants to go to a aquare dance. And inatead of Mike wanting to get rid of Paul, the farmer wanta to put the donkey in its shed.* While the subgoal structure remains the tame, the goal of wanting to put a donkey in its thed it much more telf-contained than the goal of wanting to get rid of PauL__Given thia ahift of the self-containment

*Our Mike/Paul stories were intpired by the original farmer/donkey story found in 19].

factor, we can re-examine the differing atory grammar and plot unit predictions in terms of the summary data collected for Sla and S2a. As before* 12 subjects provided summaries for each story. This time the results were dramatically different:

	SI	Sla	S2	S2a
,SUB	1.00	, .33	1.00	.58 .58

While the frequencies of the SUB units in Sla and S2a are not quite as low as the plot unit analysis might have suggested, they are significantly lower than the frequencies found for the SUB units in S1 and S2. These descrepancies are even more striking when the overall data for Sla, S2a and S3a are compared to the data for SI, S2, and S3. All three farmer/donkey stories showed extremely uniform frequency distributions for most propositions. Even so, the subgoal unit was the only non-failure-related unit to display variant behavior. This suggests that the plot unit predictions would be further substantiated under better experimental conditions.

6. Conclusions

From these experiments we conclude that summarization behavior is a function of internal memory structures along with other non-structural factors. By focusing on areas where plot unit predictions differed from story story grammar predictions, we 6aw that plot units predicted structural influences more effectively than story grammars in three out of four cases. When the structually-based plot unit predictions fell down, we saw how non-structural factors of generational subsumption and self-containment could have been operating to override structural influences. In the one case where story grammars appeared to be a clear winner over plot units, additional experimental data suggests that the originally unsuccessful plot unit prediction might have been overridden by a confounding factor of pelf-containment which no structural model can address.

All in all, the plot unit model is holding up well against experimental data, and appears to provide a promising basis for summarization algorithms.

REFERENCES

[1] Black, J. B. & Bower, G. E. (1980). Story understanding as problem solving. Poetics 9: 223-250.

[2] Dyer M. G. & Lehnert, W. G. (1981). Organization and aearch processes for narrativea. Research Report #175, Department of Computer Science, Yale University.

[3] Lehnert, Wendy, G. (1981). "Plot units and narrative summarization." Cognitive Science, (in preas)

[4] Lehnert, Wendy G. (1980). Narrative text summarization. Proceedings of The First Annual National Conference on Artificial Intelligence. Stanford University.

[5] Lehnert, W. G., Dyer, M. G., Johnson, P. N., Yang, C. J., and Harley, S. (1981). BORIS - An experiment in in-depth understanding of narrativea. Research Report #188, Department of Computer Science, Yale University.

[6] Mandler, J. M. & Johnson, N. S. (1977). Remembrance of things parsed: story structure and recall. Cognitive Psychology 9: 111-151.

[7] Reiaer, B. J., (1981). Story structure and summarization. unpubliahed manuscript. Department of Computer Science, Yale Unveraity.

[8] Rumelhart, D. E. (1975a). Notea on a schema for stories. In D.G. Bob row & A. M. Collins (Eds.), Ranreaentation and Understanding. New York: Academic Press.

[9] Rumelhart, D. E. (1975b). Understanding and summarising brief stories. In D. Laberge & S. Samuels (Eds.), Basic Procaaaing in Reading Percepation and Comprehension Hillsdale, New Jersey: Lawrence Erlbaum.

[10] Schank, R. C. (1975). Conceptual IpfprmjtJPn PrgCgmng* New York: North-Holland/American Elsevier.

[11) Thorndyke, Perry W. (1977). Cognitive structures in comprehension and memory of narrative discourse. CggPJUYC fayfihglogy. 9: 77-110.

APPENDIX A

SI: the first version of the Mike & Paul story

Mike and Paul had been cloae frienda ever since their high school days. But now Mike wanted Paul out of town for a few days so he could build a patio in Paul's backyard as a surprise birthday present. He suggested to Paul that he get away for a weekend but Paul said he wasn't interested. On another occasion Mike casually apoke about the joys of fishing or camping trips. But Paul told him he enjoyed puttering around the houae much more. Paul was getting very aettled in his old age.

Finally, Mike went to Paul's boas and aaked him to send Paul on a business trip. But Paul's boss had had a bad day and he wouldn't hear of it. Mike thought a while about what to do next. Then he had an idea.

Mike went to a friend of hia who handles the accounting recorda for Paul's company. He explained the aitation to the accountant and told him, "If you tell Paul's boas that there are irregularitiea in Paul's recorda and that you would like to examine them for a few days, the Paul will be sent away on some pretense." The accountant replied, "I'd be happy to pull the scam, but I expect a little favor in return. How about an ounce of grass?" Mike felt this was not unreasonable.

So Mike called his connection Joe and aaked him for an emergency ounce. But Joe anawered, "Sure thing, aa aoon as you pay up your tab with me." Mike personally delivered a cash payment immediately. When Joe got hia money he handed Mike an ounce. As aoon aa the accountant got the grass, he picked up the phone and called Paul's boaa. And within an hour of that phone call, Paul's boaa waa telling Paul about an emergency situation in Peoria that needed supervision. Paul was on a bua for Peoria that evening.