THE NATURE OP GENERALIZATION IN UNDERSTANDING

Michael Lebowitz

Oolunbia University, Department of Computer Science

406 Mudd Building, New York, NY 10027

ABSTRACT

True understanding of natural language text requires the inclusion of generalization and long-term memory. This paper describes the generalization • process and! memory used in the Integrated Partial Parser (IPP), a computer program that reads and remembers news stories. The need for generalization and generalization-based memory as an integral part of understanding natural language text is illustrated with examples from IPP. In addition, the nature of generalization is discussed.

1. Introduction

A computer system designed to read and remember natural language text must do more than just determine the meaning of each separate piece of text. It must also try and understand the "big picture" by comparing separate episodes. This requires making generalisations based upon the texts read that represent the general course of events in the world.

For example, proper understanding of international terrorisn requires generalizations such as those in Figure 1,

Figure 1: Typical generalizations about the world

- Terrorist attacks in Northern Ireland are carried out by members of the Irish Republican Army.
- No one is ever hurt by bombings in EL Salvador.
- The victims of kidnappings in Italy are usually businessmen.
- Takeovers in Latin America are usually carried out by left-wing groups.

The research described here was carried out at Yale University and was supported in part by the Advanced Research Projects Agency of the Department of Defense and monitored by the Office of Naval Research under contract N00014-75-C-1111.

All of these generalizations were made by a computer program, IPP (the Integrated Partial Parser), a complete understanding system that reads and remembers stories from newspapers and the UPI news wire. The generalizations made, besides being interesting observations about the world, provide useful information for text processing and the organizing memory of events. Indeed, to have a computer system that understands in as powerful fashion and over as wide a range of texts as do people it is necessary for the system to have a dynamic, generalization-based memory that can be accessed during understanding*

Further details of the topics discussed here, including the specifics of the generalization process and an extensive discussion of the memory-based parsing process developed for IPP can be found in [3] •

2. Generalization as part of understanding

As an illustration of the role of generalization in understanding, consider stories SI and S2.

(Note: All the stories used as examples in this paper are actual, unedited news stories. IPP does not require any special preparation of the stories that it reads.)

51 - Boston Globe, 5 February 79, Italy

Three gunmen kidnapped a 67 year-old retired industrialist yesterday outside his house near this north Italian town, police said.

52 - New York Times, 15 April 79, Italy

A building contractor kidnapped here on Jan. 17 was released last night after payment of an undisclosed ransom, the police said.

After reading these two stories, we know more that just the independent meanings of the stories. Even assuming no prior knowledge of terrorism in Italy, it is possible to make some tentative conclusions about the nature of kidnapping in that country. In particular, the stories lead to the plausible conclusion that businessmen are common targets for Italian kidnappers. Generalizations of this sort are a basic part of understanding.

Conclusions made so quickly are hardly sure things. However, such hypotheses allow the best information currently available to be used for future understanding and memory organization. As long as it is possible to later disconfirm these generalizations (a process discussed in [3]), then this is an effective way to make use of all existing information.

In order to generalize from a story, it must be possible to find similar examples already in memory. IFP allows the retrieval of such examples by organizing memories in terms of generalizations.

Continuing with the previous example, suppose S3, which appeared after SI and S2, is now read.

S3 - New Vbrk Times, 25 June 79, Italy

Kidnappers released an Italian shoe manufacturer here today after payment of an undisclosed ransom, the police said.

In order to fully understand this story, it is necessary to recall the previous two stories and the generalization made from it. In effect, we understand this story as being an instance of the existing generalization.

The use of generalizations shown here is a very important one - they allow a reader (human or computer) to explain events in terms of existing generalizations, rather than trying to explain each new event from scratch.

Sometimes generalizations must be used more explicitly in understanding. Fbr instance, consider the way that IPP processed story S4, after it had generalized that bombings in Spain are usually carried out by Basques (BASQUE-GEN), as shown by the computer output in Figure 2.

Figure 2: IPP inferring default role features

```
*(PARSE S4)

Story: S4 (8 24 79) SPAIN

(BOMBS EXPLODED IN A FRENCH BANK AND A FRENCH
IMMIGRATION OFFICE IN NORTHERN SPAIN EARLY TODAY
CAUSING DAMAGE BUT NO INJURIES ACCORDING TO POLICE)

>>> Beginning final memory incorporation ...

Feature analysis: EV16 (S-DESTRUCTIVE-ATTACK)
  RESULTS         AU          CAUSE-DAMAGE
  METHODS         AU          $EXPLODE-BOMB
  LOCATION        NATION      SPAIN

Indexing  EV16 as variant of BASQUE-GEN
Inferring feature ACTOR DEMAND-TYPE SEPARATISM
Inferring feature ACTOR NATION BASQUE

>>> Memory incorporation complete
```

In this example, IPP recognizes that S4 is an instance of the generalization BASQUE-GEN and uses that generalization to supply default characteristics of the terrorists. In particular, IPP assumes, corresponding with the generalization, that the terrorists are Basque separatists. Supplying information of this sort is an important use of generalizations, without which we would be required to initially provide an text processing system with every piece of information it would ever need for understanding. In this example/ the Identity of terrorists in Spain would have to be directly provided IPP if it could not learn that information itself.

3. Generalization and memory

New events are often interpreted during understanding in terms of known stereotypical situations. Fbr reasons of economy of storage and efficiency in processing, it is advantageous to record events in memory in terms of these standard situations. In order for a computer program such as IPP to begin to process text, it must begin with some knowledge of such situations, before it can begin to generalize further.

There are two types of structures that capture different forms of regularity initially provided to IPP - Simple Memory Organization Packets (S-MOPs) that describe causal stereotypes such as extortions and attacks, and Action Uhits (AUs) that represent concrete events, such as shootings, people being wounded, and hostages being released. AUs serve as modular units in the makeup of S-MOPs. I will not give a detailed description of S-MOPs or AUs in this paper, although they must be well specified, as is done in [3]. The idea of Memory Organization Packets was first introduced by Schank in [4] •

The stereotypical patterns of events captured by IPP's initial S-MOPs are not the only patterns that exist. They represent the basic information needed to understand stories. Further patterns are recognized in the form of generalizations of the sort shown earlier.

Generalized patterns serve as excellent organizers for memories of actual events, since they require only the explicit recording of unusual details (e.g. those not captured by a generalization) of a story. Furthermore, the generalizations made provide an adequate number of different points around which to organize memory.

The combination of a generalization and the events it organizes is known as a specialized MOP, or spec-MOP. As events are added to an S-MOP or spec-MOP, IPP is usually able to make generalizations that allow the memories of events to be spread among several spec-MOPs. This enables events to be stored in a distinctive, easy to retrieve fashion.

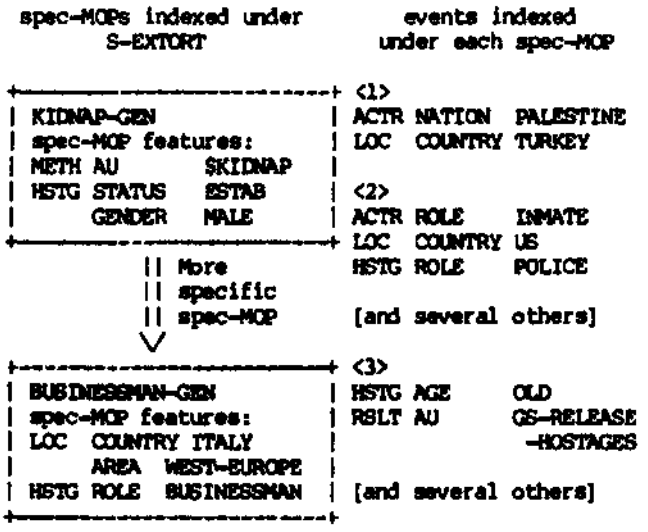
S-MOPs and spec-MOPs are fundamentally similar

structures. Presumably S-MOPs could be created from spec-MOPs that are used quite frequently. Both S-MOPs and spec-MOPs describe abstract situations. Both structures are used to organize memories of events and more specific spec-MOPs, as well as being used to make predictions for use in understanding. S-MOPs simply provide that information needed to initially understand stories.

IPP's memory, then, is a set of S-MOPs, each pointing to a net of spec-MOPs. Associated with each spec-MOP are events (in terms of the Alls and role fillers that make them up) indexed by the ways in which they differ from that spec-MOP. This index uses a discrimination net that allows easy retrieval of those events similar to a new event that might be explained by the same spec-MOP.

Figure 3 provides a concrete example of IPP's memory structure. It shows one small piece of memory, two spec-MOPs, after approximately 300 stories had been read and remembered.

Figure 3: A snapshot of IPP's memory



The section of memory in Figure 3 contains two spec-MOPs (given names for purposes of this discussion) that describe situations concerning extortion. The first spec-MOP in Figure 3, KIDNAP-GEN, describes the kidnapping of establishment males. Several events are shown indexed under it, <1> a kidnapping by Palestinians in Turkey, and <2>, a kidnapping of police by inmates in the United States.

Since each unique feature of an event is used to index it under KIDNAP-GEN, whenever this spec-MOP has been determined to be relevant to a situation it is a simple matter to find events with features we are concerned about.

A similar scheme is used to index spec-MOPs under S-MOPs and other more general spec-MOPs. Basically, all the features of the spec-MOP are used as indices pointing from the higher node to the spec-MOP. Again, this indexing simplifies the process of finding the spec-MOP at the times it is relevant.

Such indexing of spec-MOPs is shown in Figure 3. KIDNAP-GEN turns out to be a widely applicable spec-MOP, and still more specific spec-MOPs are quickly created, including one shown in Figure 3. Once IPP encounters several examples of businessmen being kidnapped in Italy, it concludes that this, too, is a generalizable situation. This decision results in the creation of a new spec-MOP, BUSINESSMAN-GEN, that is also used to organize memories of events (including <3>) •

The indexing of spec-MOPs under S-MOPs and events under spec-MOPs makes it easy to find similar events as a new story is being read. By looking at the index for spec-MOPs under the relevant S-MOP, it is possible to find generalizations that share features with an incoming story. After features have been extracted from a story, IPP uses the index for spec-MOPs under an identified S-MOP to fetch any spec-MOPs with the same features as the story.

Once IPP has found the best available spec-MOPs, it can then find events that differ from those spec-MOPs in the same ways the new story does. This time the event index is used to find other instances with some of the same non-stereotypical features•

4. The range of generalization

The inclusion of generalization in IPP lead to considerable insight about the kinds of generalizations that are useful to make as part of the understanding process. These fall into two basic classes - factual generalizations that can be made directly from information in texts or require only simple inferences and abstract generalizations that require processing of input texts beyond the initial identification of the events described. Factual generalizations are fully Implemented in the IPP as are some aspects of abstract generalization.

4.1. tactual generalization

The generalizations made by IPP are models of specific situations that are useful in future understanding. IPP concentrates on noting the common factors in similar situations. In this section I will look at the nature of factual generalization, and the basic knowledge needed to make them.

A typical factual generalization can be made from S5 and S6. IRA-GEN (paraphrased below) is the

generalization made by IPP from these two stories.

55 - Boston Globe, 12 April 79, Northern Ireland

Irish Republican Army guerrillas ambushed a military patrol in West Belfast yesterday killing one British soldier and badly wounding another army headquarters reported.

56 - New York Times, 28 August 79, Ehgland

Earl Mountbatten of Burma one of the heroes of modern British history was killed today when his fishing boat was blown up in the sea, apparently by terrorists of the Irish Republican Army.

IRA-GEN: The IRA is a common actor of terrorist attacks against the British.

In each of these stories about attacks in Great Britain, the terrorists involved belonged to the Irish Republican Army. As it indexes the two stories in memory, IPP notices that they share a number of features. From this it makes the generalization that the IRA is normally associated with terrorist attacks against British citizens.

It is important to realize that generalizations first arise from noting similarities in events, and may not be further analyzed. Ib use another terrorism example, it is quite possible to recognise that hijackings in Lebanon are usually the work of Shiite Moslems, while knowing very little about Shiite Moslems.

The factual generalizations that IPP makes tend to be observations of properties of standard role fillers in given situations and normal Action units (concrete events) for situations. Knowing detailed characteristics of potential role fillers assists in identifying actual role fillers in stories, and can supply defaults, as in Figure 2. Standard Action Uhits can supply defaults in situations in which events that take place are only partially specified.

Only a limited amount of knowledge in addition to S-MOPs and Action Units is needed to understand a sizable number of stories and begin begin to make generalizations. The properties about people, organizations, objects and countries that are initially provided IPP are shown in Figure 4.

Figure 4: Necessary properties in memory

AGE AREA AUTO BODYPART GENDER IMPORTANT

MINORITY NATIONALITY POLITICS RELIGION

ROLE STATUS TERRORIST-GROUP

Some of the information initially provided IPP to aid it in understanding later proved to be learnable from stories. For instance, the property TERRORIST-GROUP names the default terrorist organization associated with a given country. Vtien IPP was first developed, it was provided with the knowledge that the IRA normally carries out attacks in Northern Ireland, the Red Brigades in Italy and so forth.

However, it was later discovered that even if this information was deleted from IPP's initial store of knowledge, it would make similar generalizations itself. Furthermore, the knowledge from the generalizations was somewhat more specific than that Initially provided the program. For example, instead of assuming that the IRA was behind all terrorist actions in Northern Ireland, IPP was able to generalize that the IRA was often behind bombings and shootings, but not usually kidnappings and hijackings.

4.2. Abstract generalization

Not all generalizations can be made on a basic, factual level. It is also possible to make generalizations of a more abstract sort. I will consider here three classes of such generalizations —- generalizations that require high-level analysis of stories, generalizations requiring reasoning, and generalizations of other generalizations. In each case, the basic generalization process remains the same — noticing similarities among events — but the kind of information used as input is more abstract•

4.2.1. High-level generalization

It is possible to base generalizations on any of the levels of understanding that take place, not just the factual level that IPP concentrates on. All the kinds of processing that have been discussed by other researchers, such as that involving plans, goals, themes [1, 5, 8], political ACTs [6], and other high-level structures can be the sources of generalizations. The same processes — recording events in memory in terms of their representations *and* assuming that similarities among events indicate potential generalizations — works for stories analyzed at all levels.

Che example of the need for many levels of knowledge involves the need for recognizing political situations. Stories S7 and S8 are both about events that took place shortly before Rhodesia's first bi-racial elections.

S7 - UPI, 24 February 80, Fhodesia

The last day of formal campaigning for Rhodesia's independence elections was overshadowed by a powerful bomb blast Sunday that killed two people and demolished the plant of an influential Roman Catholic newspaper.

S8 - UPI, 28 February 80, Fhodesia

Gunmen reported bo be followers of guerrilla chief Robert togabe tried unsuccessfully to ambush a mobile polling station in *an* isolated attempt to wreck Rhodesia's independence election authorities reported today.

RHODESIA-GEN: Attempts at democracy in Rhodesia are opposed by terrorists.

The generalization that terrorists will try to undermine democracy in Rhodesia is one that most of us would make (or consider), but is based on rather complex analysis of these stories. In order to make such a generalization, a reader has to recognize that an attack on a newspaper undermines a free press which in turn decreases the likelihood of fair elections. Similarly, in the second story the reader must determine that an attack on a polling place will deter people from voting, which again minimizes the possibility of impartial elections.

Prom examples such as this we see that the generalization ability encompasses not only the level of concrete, factual analysis described in the previous section, but also higher-level knowledge, such as goals, plans, and political ACTS. It is necessary bo notice similarities among events at these high levels, as well as just analyzing each story to such levels.

4.2.2.. Generalization requiring reasoning

Understanding is not always a matter of simply matching up new events being described with stereotypical situations (even high-level ones) that we already know about. At times it is necessary to use sequences of inference rules to determine how a new event relates to what we already know. The results of such reasoning can be the source of generalizations in the same way as more explicit information.

Consider, for example, S9 and S10 which might lead to the generalization LEBANON-GEN.

S9 - New *York* Times, 23 January 79, Lebanon

The Palestinian guerrilla leader who reputedly planned the attack on the Israeli Olympic team at Munich in 1972 was fatally wounded here today in the explosion of a remote-controlled bomb.

S10 - UPI, 28 July 80, Lebanon

Unidentified gunmen Monday ambushed a pro-Iraqi politician riddling him and his bodyguard with bullets in the latest of the region's almost daily political assassinations.

LEBANON-GEN: Israeli agents are willing to kill their enemies in Lebanon.

Understanding these stories well enough to be able to make LEBANON-GEN requires reasoning about several different points, but I will concentrate upon the determination that both killings were carried out by Israeli agents.

Neither story mentions the identity of the actor of the killing described. In each case a rather complex chain of reasoning is needed to infer that the killings were carried out by Israeli agents. However, once these inferences have been made, it is an easy matter to make the generalization that Israelis kill their enemies in Lebanon, using the methods described in this paper. The difficult task here is to determine vdien and how to apply the relevant inference rules that make up the reasoning procedure•

Inferences such as these can easily serve as input to the generalization process. However, it is necessary to perform extensive inferencing only when needed, arri rely mostly on the stereotypical situations previously observed as much as possible, in order to avoid extensive computational problems.

4.2.3. Generalizing generalizations

The generalizations I have considered so far describe specific situations that are useful in explaining events. However, generalization is actually a multi-layer process, requiring the recognition of generalities at based upon other generalizations. Ultimately this process allows for the creation of structures applicable to wider ranges of situations.

Ib see the need for such a procedure, consider the generalizations in Figure 5.

Figure 5: Similar generalizations

ITALY-KILL-GEN: Terrorists in Italy kill people with guns.

GUAT-KILL-GEN: Terrorists in Guatemala kill people with bombs.

COL-WOUND-GEN: Terrorists in Oolombia would people with guns.

These generalizations clearly have something in conmon. Each describes a location where attacks take place, a standard result (victims being killed or wounded), and a standard method (shooting or bombing) • If we assume that features of this sort frequently appear in generalizations (i.e., if we generalize the generalizations) then we will know to expect similar features in other generalizations about attacks. This corresponds to the way we can predict elements of events from the concrete generalizations that we make.

IPP concentrates upon making accurate factual generalizations. These generalizations provide the

predictions that are needed for most of the understanding done by IPP. Some experimentation was done with high-level generalization, in particular, generalizations based on well-defined goals of terrorist groups such as black nationalists in Africa and the IRA. However, these more abstract generalizations, like those requiring reasoning and generalizations based on generalizations, are not as directly applicable to the understanding of new stories as factual generalizations and was left largely as a topic for future research.

5. Other research on generalization

Generalization, as performed by IPP, is the learning of rules that define behavior in a certain set of circumstances. While little has been done concerning learning from natural language texts, there has been work on learning in other contexts that has some relevance to IPP.

Winston's well-known program [9], is given structural descriptions of examples of blocks world constructions (and near misses) and determines the rules behind the specification of these structures. Unlike the work described here, Winston provides his system with the concepts to be learned. In addition, Winston does not have to deal with problems of organizing and finding instances in memory, since he is dealing with small numbers of cases.

A piece of work that deals with problems closer to those of IPP is that of Soloway [7]. His program takes episodes from a game of baseball, and attempts to generalize the rules of the game. The program involves multiple levels of analysis of generalizations about the data, and is able to determine the more basic rules of baseball. The most important difference here is IPP's use of generalizations to help find examples that are relevant.

Two pieces of recent research in the area of memory organization are relevant. Schank [4] introduced MOPs as a method of organizing personal episodic memories. Kolodner [2], with the program CYRUS, addresses the problems of organizing large amounts of such information in a manner suitable for efficient update, retrieval and question answering •

6. Conclusion

In this paper I have illustrated the role of generalization in the understanding process, and considered the range of generalizations that can be made. The use of generalization and memory has made IPP a powerful understanding system. IPP is written in Yale/Rutgers/UCI LISP on a DECSystem 20/60 and uses approximately 100,000 words of storage for the program (including 3200+ dictionary entries for parsing). The generalization-based memory is kept in a separate LISP core image, where over 500 events can be recorded in about 20,000 words of storage.

When IPP reads all the available stories about international terrorism taken from local newspapers and the UPI news wire, it successfully Identifies the events and role fillers for about 70-80%. Over the course of reading better than 300 stories, IPP was able to make about 125 generalizations about terrorism, approximately half of which it later rejected.

Whiile many of its generalizations were rather mundane, some were more interesting, such as the one in Figure 1 indicating that there are never casualties from bombings in El Salvador. It is generalizations of this sort that indicate the importance of including generalization in an understanding system.

REFERENCES

1. Carbonell, J. G. Jr. Subjective understanding: Computer models of belief systems. Tech. Rapt. 150, Yale University Department of Computer Science, 1979.

2. Rolodner, J. L. Retrieval and organizational strategies in conceptual memory: A computer model. Tach. Rapt. 187, Yale University Department of Computer Science, 1980.

3. Lebowitz, M. Generalization and memory in an integrated understanding system. Tech. Rept. 186, Yale University Department of Computer Science, 1980. PhD Thesis

4. Schank, R. C. "language and Memory." Cognitive Science 4, 3 (1980), 243 - 284.

5. Schank, R. C. and Abelson, R. P.. Scripts, Plans, Goals and Understanding. Lawrence Erlbaum Associates, Hillsdale, New Jersey, 1977.

6. Schank, R. C. and Carbonell, J. C. Ra: The Gettysburg Address: Representing social and political acts. In Findler, N., Ed., Associative Networks: Representation and Use of Knowledge by Computers, Academic Press, New York, 1979.

7. Soloway, E. Knowledge-directed learning. Tech. Rapt. COINS Technical Report 77-6, University of Massachusetts at Amherst, 1977.

8. Wilensky, R. Understanding goal-based stories. Tach. Rapt. 140, Yale university Department of Computer Science, 1978.

9. Winston, P. H. Learning Structural Descriptions from Examples. In P. H. Winston, Ed, The Psychology of Computer Vision, McGraw-Hill, New York, 1972.