THE USE OF EXPLANATIONS FOR SIMILARITY-BASED LEARNING*

Andrea Pohorcckyj Danyluk

Department of Computer Science - Columbia University

New York, New York 10027

ABSTRACT

Due to the difficult nature of Machine Learning, it has often been looked at in the context of "toy" domains or in more realistic domains with simplifying assumptions. We propose an integrated learning approach that combines Explanation-Based and Similarity-Based Learning methods to make learning in an inherently complex domain feasible. We discuss the use of explanations for Similarity-Based Learning and present an example from a program which applies thee ideas to the domain of terrorist events.

L Introduction

In this paper we present a novel approach to Machine Learning that integrates Explanation-Based and Similarity-Based methods to make learning in a complex domain feasible.

We consider learning by observation in the domain of acts of international terrorism. The learning mechanism itself will be a component of a system that has the following task: *Given information from a series of newspaper accounts of terrorist events, suggest an action that a law enforcement agency might take in response to those events.*

The domain of terrorist events is realistic and complex. Furthermore, since the input to the learning mechanism is limited to information from newspaper stories, we cannot assume that, in general, this information will be correct, complete, and consistent A typical description of a terrorist incident taken from the New York Times is:

> Paris, Feb 3. - An explosion, apparently caused by a bomb, ripped through a crowded shopping arcade on the Champs-Elysees tonight Eight people were wounded, three of them seriously... Witnesses said that damage was extensive.

II. System Overview

The world provides our system with descriptions of terrorist incidents in the form of newspaper articles. As we are concentrating on learning rather than natural language at this point, each newspaper account is transcribed by hand into a hierarchical frame representation. Before planning may be done in response to an incident, the incident description passes through a learning and analysis module. This module has access to a rule base, a hierarchy of general concepts, and a hierarchy of previously categorized terrorist events. After a new event is analyzed, it is placed in the incident hierarchy.

The learning and analysis module provides the planner with a generalized version of the terrorist incident as well as a detailed explanation of the event At this time, the learning and analysis module (Figure 1) is the focus of our research and will be described in the remainder of this paper.

*This research was supported in part by the Defense Advanced Research Projects Agency under contract N00039-84-C-0165. Discussions with Michael Lebowitz were quite useful as were his comments on earlier drafts of this paper.

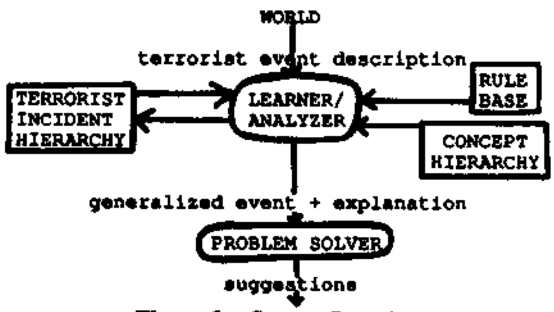

Figure 1: System Overview

III. Two Learning Methods

Two Machine Learning methods have received a great deal of attention recently: Similarity-Based Learning and Explanation-Based Learning.

Similarity-Based Learning (SBL) (e.g., [Winston 72; Michalski 83; Lebowitz 86a]) involves the comparison of several instances of a concept in order to find features shared among them and differences between them. Common features are assumed to define a useful concept.

Explanation-Based Learning (EBL) (e.g., [DeJong 83; Mitchell 83; Mitchell 86]) involves a knowledge-intensive analysis of a single instance of a concept The learner attempts to explain the instance and then creates a general concept consistent with the explanation.

Both SBL and EBL have been shown to adequately handle learning in a variety of domains. However, SBL methods are often influenced by coincidence. Furthermore, one could argue that a system cannot gain any real understanding by simple feature comparison. These shortcomings are not shared by EBL methods. However, the generation of *detailed* causal explanations is problematic if our domain model is neither complete nor necessarily consistent and becomes computationally infeasible as the complexity of the problem domain increases.

IV. Combining EBL and SBL

We have found that, despite their respective weaknesses, EBL and SBL may be *combined* to make learning in a complex domain feasible. Figure 2 describes some of the information exchange that will be necessary in a system integrating the two methods.

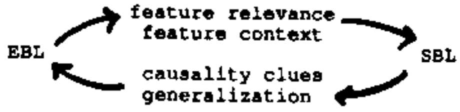

Figure 2: Combining EBL and SBL

An explanation of a terrorist event provides clues as to the features of the event that might be relevant to it If a feature of the event actually plays a role in the explanation, we can consider it more "important" than those that do not. When attempting to match instances for common features, the SBL module can constrain its search to those features that the EBL module finds to be "more important".

The EBL module also provides the SBL module with contextual clues for matching. A feature of an event appears, not in isolation, but with other features. The way a feature value should be viewed may be determined from the context in which it appears (i.e., considering the other features that exist.)

The SBL module, in turn, provides information to the EBL module. It provides causality clues that help the EBL module choose rules to explain an event [Lebowitz 86b]. It also controls the range of hypothetical generalizations of an event that an EBL module might consider.

The research described here differs from previous work in Integrated Learning [Lebowitz 86b; Pazzani 85] in a number of ways. Although we rely on SBL to guide the explanation process as Lebowitz does, our SBL phase is first provided with a surface understanding of the events under consideration. As a result, our SBL phase is not as succeptible to the influence of coincidence. Pazzani makes generalizations based upon justification provided by an explanation. Recognition of a similarity between instances of a concept may cause an explanation to be attempted in order to justify the similarity. Coincidental similarities may result in unnecessary explanatory work. Finally, our SBL phase allows for inexact match, which neither of the previously mentioned systems does.

In the remainder of this paper we will discuss the upper half of the cycle shown in Figure 2. We will show our system's simple Explanation-Based (EB) phase and then how it is used by the Similarity-Based (SB) phase.

V. The Explanation-Based Phase

In this section we describe the initial EB phase of our system that analyzes a terrorist event. This phase borrows many ideas from other EBL models (especially [DeJong 83]). It has access to a knowledge base of general rules about terrorism. This initial phase uses an abridged version of the set used by subsequent EB phases. To derive an actual explanation (causal description) of a terrorist incident, it uses backchaining. We have chosen the following explanation goal for this domain: *Explain why this incident is a terrorist event.*

Consider the newspaper story given in Section I. The system is given a hierarchical frame-based representation of this incident. It uses a set of (possibly inconsistent) rules to develop a simple causal explanation of those factors which cause the incident to be considered a terrorist event The explanation generated by the system is shown in Figure 3.

```
(back-chain 'event1 '(fact terrorist_event))

(FACT EXPLOSION)<-((HAS CAUSE1 INSTRUMENT BOMB)
        (NOT-HAS CAUSE1 REVERSE-ACT DEFUSED))
(FACT CROWDED_TARGET)
            <-((GEN LOCATION1 TYPE PUBLIC_PLACE)
                (HAS LOCATION1 AVAILABLE OPEN)
                (NOT (FACT SIGNIFICANT_WARNING)))
(FACT HURT_PEOPLE)<-((FACT EXPLOSION)
                    (FACT CROWDED_TARGET)
                (NOT (FACT SIGNIFICANT_WARNING)))
(FACT PURPOSE_ACHIEVD)<-((FACT HURT_PEOPLE))
(FACT PROPERTY_DESTROYED)
            <-((HAS RESULT1 DAMAGE EXTENSIVE))
(FACT PURPOSE_ACHIEVD)
                  <-((FACT PROPERTY_DESTROYED))
(FACT SUCCESSFUL_ACT)<-((FACT PURPOSE_ACHIEVD))
(FACT TERRORIST_EVENT)<-((FACT SUCCESSFUL_ACT))
```

Figure 3: System-Generated Explanation

We claim that finding a simple causal explanation of an event is computationally feasible and, despite its simplicity, useful in providing information for the SB phase of the system.

The simple explanation provides clues as to the relative importance of features of the incident. In the above example, the EB phase used information about the location of the event in explaining it. As a result, the SB phase concentrates on this feature.

The explanation also provides clues as to the contexts in which some of the features might be viewed. For example, the fact that the incident occurred in a public place at a time when it was open to the public indicates that the hour of the event's occurrence should be viewed as a time when public places are generally open, rather than, say, as a time when a given television show is on. When considering other events, the SB phase considers their times of occurrence in the same context.

VI. The Similarity-Based Phase

Once we have a simple explanation of an event, we compare the event to incidents previously seen by the system. The comparison provides us with a hierarchical frame structure representing a more general terrorist event scenario.

Earlier work on SBL in the terrorism domain was done by Lebowitz [Lebowitz 83; Lebowitz 86a]. In that work, however, only exact matches on features were allowed. The quality of match between two incidents was determined only by the number of features they had in common.

We have improved upon this earlier work in a number of ways. We represent events as hierarchically organized sets of frames rather than as simple feature lists. Our system allows for inexact matches between feature values and estimates the quality of the match. The match takes into account the explanation of the event by using the information about feature relevance and context determined earlier by the EB phase.

To illustrate these points we consider the match of the event described above and the following incident as described in the New York Times:

> Paris, Feb. 4 - A bomb ripped through a crowded bookstore in the Latin Quarter tonight, wounding four people. The bomb was planted in the basement record section of the Gibert Jeune bookstore on the Place St-Michel. The blast occurred at 7:40 P.M. with scores of customers in the store.

An abridged version of the representations of these two events as given to the system appear in Figure 4. The representation of the first event includes information about feature importance and context provided earlier by the EB phase as indicated by *EFACT* and *CTEXT*, respectively.

```
EVENT1:                                  EVENT2:
CAUSE: CAUSE1                            CAUSE: CAUSE2
 INSTRUMENT: BOMB   *EFACT*               INSTRUMENT: BOMB
 REVERSE-ACT: NONE   *EFACT*              REVERSE-ACT: NONE
LOCATION: LOCATION1                      LOCATION: LOCATION2
 CITY: PARIS                              CITY: PARIS
 PLACE: PLACE1                           PLACE: PLACE2
  ENVIRON: PUBLIC                          ENVIRON: PUBLIC
  AVAILABLE: OPEN   *EFACT*                AVAILABLE: OPEN
  TYPE: SHOP_ARCADE   *EFACT*             TYPE: STORE2
                                            TYPE: BOOK_STORE
                                            SECTION: RECORD
  NAME: NIL                                NAME: GIBERT_JEUNE
  ADDR: CHAMPS_ELYSEES                     ADDR: PL_ST_MICHEL
TIME: TIME1                               TIME: TIME2
 HR: 21.00   *CTEXT=BUSINESS*             HR: 19.40
RESULT: RESULT1                           RESULT: RESULT2
 NUM_KILLED: 0                             NUM_KILLED: 0
 DAMAGE: EXTENSIVE   *EFACT*               DAMAGE: EXTENSIVE
WARNING: NONE   *EFACT*                   WARNING: NONE
```

Figure 4: System Input

The SB matcher docs frame-based matching. First an exact match on corresponding slot values is attempted.

In addition to exact matches we would like the system to recognize that two different values might be similar enough to constitute a match. For example, the first incident took place in a shopping arcade while the second took place in a store. A comparison of these events based on exact matching would consider these places different. Our matcher looks for some link between the values. Recall from Section II that the matcher has access to a hierarchy of general concepts. Figure 5 shows a fragment of the hierarchy. The matcher, given two values, climbs the hierarchy to find a general concept that classifies both. Given the values STORE and SHOP_ARCADE, the system finds that they are both SHOP AREAs. The system also gives a measure of how similar tEe values are.

Figure 5: Fragment of the Concept Hierarchy

The concept hierarchy is not strictly a tree, so there may be more than one generalization under which a concept falls. Here the context information provided by the EB phase is most valuable. It is used to constrain which parents the matcher looks at In the absence of context information, the system chooses the generalization with smallest matching distance.

Figure 6 shows the frame hierarchy for the composite incident based upon a match of the two events described above. DIST specifies the quality of match, 0 indicating exact match. This structure is used to build a generalization-based memory [Lebowitz 86c].

To summarize, the SB matching phase does a frame-based match using general information about the world and context clues provided by the earlier EB phase. It creates a composite event scenario based on an understanding of the events considered.

VII. Conclusion

We have described a program that uses a simple explanation to guide SBL in the domain of international terrorist events. Applying earlier Machine Learning techniques to a complex domain such as this one introduces new problems. It becomes feasible to look at such domains by integrating EBL and SBL. Simple explanations focus the attention of the SB matcher which will in turn provide causality clues that make the derivation of complex explanations feasible.

```
(match '(event1 event2))
COMPARING FRAMES:(EVENT1 EVENT2)
EVENT1_EVENT2:
CAUSE: CAUSE1_CAUSE2
  INSTRUMENT: BOMB   DIST: 0
  REVERSE-ACT: NONE   DIST: 0
LOCATION: LOCATION1_LOCATION2
  CITY: PARIS   DIST: 0
  PLACE: PLACE1_PLACE2
    ENVIRON: PUBLIC   DIST: 0
    AVAILABLE: OPEN   DIST: 0
    TYPE: SHOP_AREA   DIST: 2
    NAME: NIL   DIST: NIL
    ADDR: CENTRAL_PARIS   DIST: 3
TIME: TIME1_TIME2
  HR: BUSINESS_HOURS   DIST: 2
RESULT: RESULT1_RESULT2
  NUM_KILLED: 0   DIST: 0
  DAMAGE: EXTENSIVE   DIST: 0
WARNING: NONE   DIST: 0
```

Figure 6: Composite Event

References

[Dejong 83] DeJong, G. F. An Approach to Learning from Observation. Proc. First International Machine Learning Workshop, Champaign-Urbana, Dlinois, 1983.

[Lebowitz 83] Lebowitz, M. "Generalization from Natural Language Text." *Cog. Sci.* 7,1983.

[Lebowitz 86a] Lebowitz, M. Unimem, A General Learning System: An Overview. Proc. ECAI-86, Brighton, England, 1986.

[Lebowitz 86b] Lebowitz, M. "Integrated Learning: Controlling Explanation." *Cog. Sci. 10,*1986.

[Lebowitz 86c] Lebowitz, M. Concept Learning in a Rich Input Domain: Generalization-Based Memory. In R. S. Michalski, J. G. Carbonell, and T. M. Mitchell, Ed., *Machine Learning: An Artificial Intelligence Approach, Volume II,* Morgan Kaufmann, Los Altos, California, 1986.

[Michalski 83] Michalski, R. S. "A Theory and Methodology of Inductive Learning." *AI 20,* 1983.

[Mitchell 83] Mitchell, T. Learning and Problem Solving. Proc. UCAI-83, Karlsruhe, West Germany, 1983.

[Mitchell 86] Mitchell, T., Keller, R., and Kedar-Cabelli, S. * 'Explanation-Based Generalization: A Unifying View." *Machine Learning 1,* 1986.

[Pazzan85] Pazzani, M. Explanation and Generalization Based Memory. Technical Report UCLA-AI-85-13, UCLA, 1985.

[Winston 72] Winston, P. H. Learning Structural Descriptions from Examples. In P. H. Winston, Ed., *The Psychology of Computer Vision,* McGraw-Hill, New York, 1972.