ON THE NATURE OF EXPLANATION OR WHY DID THE WINE BOTTLE SHATTER

by

R. S. Michalski
H. Ko

Present Workshop on Explanation-based Learning, Stanford Univ. Mar. 1988.

ON THE NATURE OF EXPLANATION
or
WHY DID THE WINE BOTTLE SHATTER?[1]

R. S. Michalski
George Mason University

H. Ko
University of Illinois at Urbana-Champaign

INTRODUCTION

An explanation of some phenomenon involves determining knowledge that logically entails the phenomenon, and creating a structure that demonstrates that this knowledge indeed entails the phenomenon. This distinction is useful in gaining uniform perspective of various learning methodolgies and leads us to a formulation of a new learning paradigm unifying inductive and deductive learning.

To illustrate this distinction, suppose a person did not come to a meeting with a friend. He may later told his friend that his mother got sick. This statement, called *explanation knowledge*, explains the person's behavior to his friend, if the friend has the background knowledge that a crisis in a family typically overrides other commitments, and that a mother's sickness is a family crisis. The explicit demonstration that the explanation knowledge indeed explains the behavior (which in our example case is a simple two step application of modus ponens) is called *explanation structure.* The explanation structure is thus a proof that the explanation knowledge together with background knowledge logically implies the behavior, or in general, any phenomenon that is being explained.

In analytic learning, such as explanation-based learning, without knowing the illness of the person's mother, no explanation can be constructed. People in such cases, however, usually construct some hypothesis to explain the absence at a meeting. The friend, for example, may hypothesize that something serious might have happened to the person. This leap of faith is a form of constructive inductive inference [Michalski, 1983]. It is done by employing the background knowledge that if a person, who normally comes to meetings, did not come, then it is likely that something serious is the reason.

To handle such problems, *constructive closed-loop learning* (or, briefly, *constructive learning*) uses deductive inference when the observation is logically implied by background knowledge, but resorts to inductive inference when an additional (or modified) knowledge is necessary to establish the needed explanation structure. Thus, such a system can learn when its initial knowledge is inadequate.

TYPES OF EXPLANATION

To explain some observation to a person means to construct a knowledge structure that conceptually relates the person's background knowledge (BK) with the observation statement (OS). Formally, the constructed knowledge structure, called *explanation structure* (ES), must

[1]This research was supported in part by the Office of Naval Research under grant No. N00014-82-K-0186 and in part by the Defense Advanced Project Agency under the grant administered by the Office of Naval Research No. N00014-K-85-0878

demonstrate that BK logically entails OS, which we write:

In other words, OS must be a logical consequence of BK. In many situations, however, the background knowledge (BK) may not be adequate to establish (1). BK may be inadequate because it may be insufficient, intractable (too complex), or inconsistent with the observations. In all such cases, BK has to be modified or enhanced by additional knowledge, called *explanation knowledge* (EK). The explanation knowledge may be given to us from another source, e.g., teacher or environment, or may have to be hypothesized. In these cases, the *explanation structure* (ES) is a proof that

c cw(3i) r. BK |> OS (1)

Constructing an explanation knowledge may require changing (updating, correcting, etc.) BK into some modified BK*. Thus, in general, (2) is in the form

c cw(3i) r. BK & EK |> OS (2)

Based on these considerations, we can say that an explanation of an observation consists of components:

- EK - explanation knowledge,
- ES - an explanation structure which demonstrates that explanation knowledge together with background knowledge logically entails observation.

Using this conceptual framework, we can unify different methods of learning. In explanation-based learning, EK is null, and one seeks the *explanation structure* (ES) that shows (1). In empirical inductive learning, BK is small and inadequate for explaining OS, so one needs to hypothesize explanation knowledge (EK). This is done by drawing inductive inference from input observations. There may be many different situations between the two extremes, the purely empirical and purely analytic learning. When BK is inadequate, and one needs either to create EK so that (2) holds, and/or modify BK so that (3) holds. In constructive induction, BK may be substantial, but still inadequate for deductively explaining the observation [Michalski 1983]. In general, constructive learning describes a general paradigm intended to handle all situations.

The above leads us to the distinction between two types of explanation:

- **a deductive explanation** which consists merely of the explanation structure demonstrating that the observation, OS, is a logical consequence of what the system already knows (BK[2]), i.e., BK |> OS. The explanation knowledge is null.
- **an inductive explanation** which consists of an explanation knowledge (EK) that is inductively hypothesized, and the explanation structure that demonstrates EK together with (possibly modified) background knowledge, BK*, and the inductive explanation implies the observation (OS), i.e., BK* & EK |> OS.

Let us characterize the empirical inductive learning, analytic learning, and constructive learning methodologies in terms of introduced concepts.

[2]In explanation based learning literature, BK is typically called domain knowledge. In general, BK contains both domain knowledge and general meta knowledge including control knowledge, rules of inference and so on.

We will illustrate the methodologies by an example. This example deals with explaining why a wine bottle put in a refrigerator shattered. We can characterize by the following observational statements.

Observational Statements:

1. The wine is contained in the bottle.
2. The wine bottle is left in a refrigerator.
3. The wine bottle is shattered.

Let us consider different scenarios for constructing an explanation structure to answer why the wine bottle was shattered.

EMPIRICAL INDUCTIVE LEARNING

The work on empirical inductive learning (EIL) presupposes little background knowledge relevant to the task at hand, so that the main concern is to hypothesize a concept or rule primarily on the basis of the observational data supplied to the system [Michalski, 1983, 1987]. Since there is usually a plethora of possible hypotheses that would possibly explain the observation, the main problem is to find the most plausible, or generally, the most preferred explanation. Thus, the main inference scheme of these systems is inductive. We can describe the learning system.

Given:

- *Observational statements* (OS) about an object, phenomenon, or a process.
- *Background knowledge* (BK) which includes domain concepts. the preference criterion for choosing among competing hypotheses. and inductive rules of inference.

Determine:

- *Explanation knowledge* (EK) that. if true, logically entails the observation and is most plausible. or. in general. most desirable among all other such hypotheses according to a given preference criterion.

Explanation structure (ES) in most EIL systems involves subsumption relationships between EK and OS. Thus, they use matching procedures to establish ES. Suppose only domain specific descriptors about objects in the observation are known in the wine bottle example.

Background Knowledge:

4. A wine is a liquid.
5. A bottle is a rigid object.

EIL system would postulate a surface *explanation knowledge* that the wine bottle was shattered because, e.g.

- it always breaks in a refrigerator no matter what it contains. or
- it contains a liquid.

There are many more possible EK's. However, there is little guidance as to which is more plausible other than some syntactic preference criteria. So, these systems resort to multiple observations of similar phenomena. In this case. the learner may encounter a bottle of milk in the refrigerator and later observe that the bottle was broken – then, it may prefer the second hypothesis.

ANALYTIC LEARNING

Most known forms of analytic learning are explanation-based generalization [Mitchell & Keller & Kedar-Cabelli 1986] and explanation-based learning [DeJong & Mooney, 1986]. In this approach, the system attempts to show that the background knowledge of the observer accounts for the observation. A successful explanation enables the system to formulate a more efficient or operational rule to account for the observation. The basic inference scheme used in these systems is deductive.

Given:

- Observational statements (OS) about some objects, phenomena. and processes.
- Background knowledge (BK) which contains general and domain-specific concepts for interpreting the observations and inference rules relevant to the domain and deductive rules of inference.

Determine:

- A reformulation of the background knowledge that logically entails the observation and is more effective and/or efficient than the prior knowledge.

The explanation structure (ES) in these systems is either a proof tree generated by a theorem prover, or some other equivalent form. In the wine example. the observer may possess more knowledge in addition to those of EIL system.

Background Knowledge:

6. Refrigerator freezes the objects inside.
7. When a liquid is frozen, its volume increases.
8. When a rigid object is frozen. its volume decreases.
9. When the volume of a contained object is bigger than than that of the container, the container will break.

Explanation Structure:

10. The wine is frozen because 1. 2. and 6 are believed.
11. The volume of the wine increases because 4. 7. and. 10 are believed.
12. The volume of the wine bottle decreased because 2, 5. 6 and 8 are believed.

Now. one can conclude that the wine bottle is shattered from 9. 11 and 12. In EBL system. a reformulation of background knowledge is sought that would make it more effective and/or efficient.

- When one leaves a rigid container full of liquid in a refrigerator. it will break.

CONSTRUCTIVE LEARNING

Constructive induction [Michalski 1983] incorporates domain specific knowledge and a deductive reasoning mechanism into an empirical inductive learning system. The learning system uses this knowledge to construct new descriptors and concepts not present in the observational statements. in order to generate more adequate description space for inductive hypotheses. Recently, the concept of constructive induction was generalized to constructive learning.

Constructive learning integrates inductive learning (empirical and constructive) with deductive (analytical) learning, and includes also the ability to determine the task relevant knowledge in the knowledge base, and to evaluate the constructed knowledge in order to decide if it is to be stored in the knowledge base for future use. A CL system is able to use full deductive decision procedure to establish explanation structure, as well as to synthesize plausible explanation knowledge, like empirical inductive learning systems.

In the wine example, a CL system may have additional background knowledge than an EIL system but less than an EBL system:

Background Knowledge

6. When the volume of a contained object is bigger than that of the container, the container will break.

Note, it does not know the relationships between the volume of objects and their frozen states.

In CL, the system may generate several different hypotheses (explanation knowledge) that entail the observation.

1. A glass breaks in low temperature.
2. A glass shrinks in low temperature.
3. The volume of wine inside the bottle expands in low temperature.

Hypothesis 1 may be dismissed after one recalls an empty glass in the refrigerator before and it did not break. Hypotheses 2 and 3 is not contradicted by past memory but it is difficult One may construct an experiment. Leave a container full of ice in the refrigerator and it did not break. Then. hypothesis 2 can be dismissed because if true. the container in the experiment should have been broken. Then, hypothesis 3 may be strengthened with an additional experiment with different liquids inside and it may conclude any liquid inside a bottle will break in low temperature.

CONCLUSION

Explanation is consists of two components: the structure that was built and the new knowledge that were needed to form the structure. EBL and EIL systems are compared based on this distinction. The constructive learning, which integrates these learning paradigms within a single conceptual framework, is an important new research direction for machine learning.

REFERENCES

(1) Michalski, R. S., "Theory and Methodology of Inductive Learning," in *Machine Learning: An Artificial Intelligence Approach*, R. S. Michalski, J. G. Carbonell. T. M. Mitchell (eds.). Tioga Publishing Co., 1983.

(2) Mitchell, T. M., Keller, T., and Kedar-Cabelli, S., "Explanation-Based Generalization: A Unifying View," in *Machine Learning Journal*, vol 1, January 1986.

(3) DeJong, G. and Mooney, R., "Explanation-Based Learning: An Alternative View." in *Machine Learning Journal*, vol 2, 1986.

(4) Michalski, R. S., "Concept Learning," in the Encyclopedia of Artificial Intelligence. John Wiley & Sons. E. Shapiro (eds), January 1987.