Title
Belief Revision and Induction

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 9(0)

Authors
Rosenberg, Donald
Langley, Pat

Publication Date
1987

Peer reviewed

Belief Revision and Induction

Donald Rose
Pat Langley

Irvine Computational Intelligence Project
Department of Information & Computer Science
University of California, Irvine CA 92717 USA

Abstract

This paper describes how inductively produced generalizations can influence the process of belief revision, drawing examples from a computational model of scientific discovery called REVOLVER. This system constructs componential models in chemistry, using techniques from truth maintenance systems to resolve inconsistencies that arise in the course of model formulation. The latter process involves reinterpreting observations (premises) given to the system and selecting the best of several plausible revisions to make. We will see how generalizations aid in such decisions. The choice is made by considering three main factors: the number of models each premise supports, the number of premises supporting the generalized reaction, and whether a proposed revision to that premise matches any predictions made by any generalizations. Based on these factors, a cost is assigned to each premise being considered for revision; the hypothesis (set of revisions) having the lowest cost is chosen as best, and its revisions are carried out. By viewing generalized premise reactions as a paradigm, we will argue that the revision process of REVOLVER models how scientific paradigms shift over time.

Introduction

In this paper, we discuss three main topics: how to form simple scientific theories, how to revise theories in order to account for new information, and how empirical generalizations can help to direct that revision process. In earlier papers, we have discussed STAHLp (Rose & Langley, 1986a, 1986b), a computer program that discovers explanatory models of chemical substances. In this respect it was similar to the STAHL system (Zytkow & Simon, 1986), which also constructed such models. However, unlike its predecessor, STAHLp featured a unified mechanism for reinterpreting its given observations when inconsistencies arose.

Yet STAHLp itself had a number of limitations. One of the most important is the need to take generalizations into account during the belief revision process. For instance, one should recognize when revising a premise will affect strongly held generalizations – i.e., those supported by many observations. Taken together, a set of generalizations can be viewed as a paradigm in the sense of Kuhn (1970). For example, many of STAHLp's runs involved observations associated with two main paradigms of 18th century chemistry: phlogiston theory and oxygen theory. The phlogiston framework, which came first historically, was based on the assumption that burning substances emitted a substance (phlogiston) during combustion. Oxygen theory took an opposing view of this process, stating that when a substance burns, it gains another substance (oxygen) in the process.

Historically, scientists like Lavoisier used generalizations to argue for their paradigm (e.g., oxygen theory) and to reinterpret observations made by supporters of competing paradigms (e.g., phlogiston theory). Gradually, predictions made from the general reactions summarizing the oxygen paradigm were confirmed by new experiments, many of which were proposed after the observations of the phlogiston paradigm were reinterpreted. Later in the paper, we describe REVOLVER, a model of scientific theory formation that takes such inductive generalizations into account during its belief revision process. But first, let us recount the earlier work on STAHLp.

STAHLp: Scientific Discovery and Belief Revision

As we have mentioned, STAHLp constructed componential models based on many kinds of observations. For example, suppose the system is given initial beliefs from phlogiston theory: charcoal and calx-of-iron (known as iron oxide today) react to form iron and ash, and charcoal decomposes into phlogiston and ash. In shorter notation, the premises are $\{CI\ Ch\} \rightarrow \{I\ Ash\}$ and $\{Ch\} \rightarrow \{Ph\ Ash\}$. The program would first infer the components of charcoal ($Ch = \{Ph\ Ash\}$), then substitute its components into the first reaction, yielding $\{CI\ Ph\ Ash\} \rightarrow \{I\ Ash\}$. Cancelling ash from both sides yields $\{CI\ Ph\} \rightarrow \{I\}$, and the system would now infer a model for iron ($I = \{CI\ Ph\}$).

Using this method, STAHLp constructed many componential models, replicating several episodes from the history of science. However, in the process of discovering such componential models, inconsistencies can arise. This occurs when the premises leading to these beliefs are themselves mutually inconsistent; either specific observations may be faulty, or groups of premises cannot be believed simultaneously. In both cases, some observations must be reinterpreted in order to arrive at a consensus. In an attempt to model how scientists reinterpret their observations when confronted with inconsistencies, STAHLp used belief revision techniques based on those of truth maintenance systems (Doyle, 1979; de Kleer, 1984).

Let us look at an example of how STAHLp handles the task of reinterpreting (i.e., revising) its premises. The first premise given to the system is again from phlogiston theory: the belief that mercury decomposes into calx-of-mercury and phlogiston ($\{M\} \rightarrow \{CM\ Ph\}$). The model $M = \{CM\ Ph\}$ is then inferred. Next STAHLp is given a second premise, one which embodies oxygen theory: $\{M\ O\} \rightarrow \{CM\}$. Substituting mercury's components into the second premise yields $\{CM\ Ph\ O\} \rightarrow \{CM\}$, and cancelling CM from both sides of this transformed reaction results in $\{Ph\ O\} \rightarrow \{\}$. This is an inconsistent reaction, because it has inputs but no outputs.

At this point STAHLp invokes belief revision to find the premises that caused this error, propose revisions to those premises, and implement the best set of revisions. Each proposed revision to a premise can be viewed as a reinterpretation of the observation encapsulated by that premise. In our example, the system proposes four sets of revisions (hypotheses), each of which would remove the inconsistent reaction:

(1) Premise 2: outputs really had Ph and O;
(2) Premise 2: outputs really had Ph, inputs really had no O;
(3) Premise 1: inputs really had O, outputs really had no Ph;
(4) Premise 1: outputs really had no Ph; Premise 2: inputs really had no O.

Now the system must evaluate each hypothesis. STAHLp used one heuristic to drive its evaluation: *prefer the revision of premises that support the least number of models.* That is, the system tried to reach a consensus by altering the current theory (set of models plus the inconsistency) in the least drastic way. In our example, the cost of revising premise 1 is **2** because that premise supports the model of mercury and the inconsistency. In contrast, premise 2 has a cost of **1** because it supports only the inconsistency. Since the cost of each hypothesis is the cost of its suggested revisions, the four hypotheses have a cost of **1**, **1**, **2** and **2**, respectively. The system then selects the best (lowest cost) hypothesis – in this case, either the first or second hypothesis. Either set of revisions results in removal of the inconsistency when inferencing begins again, and thus the premises will be mutually consistent.

REVOLVER: Using Generalizations to Influence Belief Revision

We have seen that the STAHLp program modelled an important aspect of scientific discovery: the need to reinterpret one's observations when conflicts cause an inconsistent theory to be formed. While STAHLp's successes were significant, its belief revision process left no place for generalizations like those used by Lavoisier. In response to this limitation, we are integrating inductive reasoning into REVOLVER, a new model of scientific theory formation. This system will be able to use generalizations as part of the belief revision process and, ultimately, to formulate these generalizations on its own initiative.

The use of generalization in REVOLVER takes the form of two new heuristics, incorporated into the evaluation function that decides which revisions to make during belief revision. New heuristic (1) is used to *prefer revision to premises that support relatively weak generalized beliefs*. For example, when considering which of a set of premises to revise, REVOLVER would change the premise that led to the generalization having the least number of supporting premises, all other factors being equal. When selecting among candidate revisions, new heuristic (2) is used to *prefer revisions that confirm predictions made by strong generalized beliefs*. For example, if plausible revisions have been generated, and only one of them matches a prediction made by some generalization, then REVOLVER would select that revision, all other factors being equal. If there are several such matches, the system would select the revision matching the prediction that is part of the most heavily supported generalization.

Let us take a closer look at how these rules will be incorporated into REVOLVER. The method of detecting inconsistencies and generating plausible revisions will remain the same; only the function used to evaluate the revisions will change. The new evaluation function selects revisions of premises which support few models and generalizations, and which match a prediction if possible. Ignoring prediction matching for the moment, the new evaluation function computes the cost for each proposed revision by adding the number of models supported by the premise to be revised, plus the *total* number of premises supporting each generalization to which that premise lends support. The higher this number for a given premise, the more damage would be done to the belief system by revising that premise. However, if revising a premise would match a prediction made by a generalization, then the cost would be decreased, indicating that the revision is more desirable.

To illustrate how these new rules will be used in REVOLVER, let us reanalyze our previous example, taking into account the above changes. The previous example only involved two premises: $\{M\} \rightarrow \{CM\ Ph\}$ (premise 1) and $\{M\ O\} \rightarrow \{CM\}$ (premise 2).

Suppose two new premises are added to the system: {I} $\rightarrow$ {CI Ph} (premise 3) and {I O} $\rightarrow$ {CI} (premise 4), where I represents iron and CI represents calx-of-iron. At this point, the system can form two general reactions: premises 1 and 3 lead to {X} $\rightarrow$ {CX Ph}, while premises 2 and 4 lead to {Y O} $\rightarrow$ {CY}. The new classes X and Y represent those substances obeying the general reactions described (in this example, both classes contain M and I). The first generalization represents phlogiston theory: when any combustible burns, the calx of that substance (an element) remains and phlogiston is emitted. The second generalization represents oxygen theory: when any combustible burns, oxygen is gained and the calx of that substance (a compound) remains.

After this inductive step ends, standard deduction takes place. REVOLVER follows the same inference path as before, arriving at M = {CM Ph} and then at the inconsistent reaction {Ph O} $\rightarrow$ {}. The sets of revisions (hypotheses) proposed are also the same. However, the cost assigned to each is now different. Let us reexamine the four hypotheses proposed by both STAHLp and REVOLVER, along with the costs assigned by each system. Each of the first two hypotheses involved changing only premise 2 ({M O} $\rightarrow$ {CM}). While STAHLp would assign each a cost of 1 (because premise 2 only supports the inconsistency), the new REVOLVER would assign each hypothesis a cost of **3**, since premise 2 also supports the generalization {X O} $\rightarrow$ {CX}, which has two premises as support. The third hypothesis involves changing premise 1 ({M} $\rightarrow$ {CM Ph}). While STAHLp would assign a cost of 2 (because premise 1 supports one model plus the inconsistency), REVOLVER would assign a cost of **4**, since premise 1 also supports the generalization {X} $\rightarrow$ {CX Ph}, which has two premises as support. The fourth hypothesis involves changing both premise 1 and premise 2. STAHLp would again assign a cost of 2, since premise 1 and 2 together support one model plus the inconsistency; in contrast, REVOLVER would assign a cost of **6**, since each premise supports a generalization that has two premises as support.

To summarize, STAHLp's hypotheses had costs of 1, 1, 2 and 2, respectively; those of REVOLVER had costs of **3**, **3**, **4** and **6**. Note that the first two hypotheses will be considered the best by both systems, but that the last hypothesis is clearly the worst in the view of REVOLVER, since it involves premise changes that would affect two generalizations (all others would impact only one generalization). If we continue altering this example by adding more premises, the choice of best hypothesis will also become different between the two systems. In particular, consider the addition of another premise that fits oxygen theory. This would mean that its generalization ({X O} $\rightarrow$ {CX}) would now have three supporting premises, and thus the hypothesis costs would now become **4**, **4**, **4** and **7**, respectively. Note that three hypotheses now tie for best; the new entry is the third hypothesis, which suggests revising only premise 1 – a belief from phlogiston theory.

In other words, adding more support to the oxygen theory generalization makes revision of the phlogiston theory premise more plausible. This trend continues further if we add yet another oxygen theory premise; this increases the support of its associated generalization to four premises. The new hypothesis costs would thus be **5**, **5**, **4** and **8**, respectively; note that the third hypothesis is now the sole best choice. In short, as the general belief embodying oxygen theory gained in strength, it became less desirable to revise premise 2 (which supports it), and thus more desirable to revise premise 1 (which supports the general belief embodying phlogiston theory).

Discussion

While the previous example involved induction on a relatively small number of premises, one can envision examples involving large systems of beliefs, where the generalizations embody laws supported by substantial observational evidence. If we look at each law (plus its supporting premises) as a paradigm and note that competing paradigms may result from different subsets of the premises, we feel that our new heuristics relating induction and belief revision bring us closer to modelling how paradigms shift over time. In this view, older paradigms would usually consist of well-supported generalizations (i.e., laws that summarize many observations), while new paradigms would consist of generalizations having only a few observations as supporting evidence. Initially, new heuristic (1) – preferring revisions to premises supporting weak generalizations – would tend to protect older paradigms. We claim that this is a plausible model of how science normally proceeds; old paradigms tend to become entrenched and require a steady accumulation of negative evidence to overthrow them.

This negative evidence comes as more observations are gathered that fit the generalization for the competing paradigm. In this manner, the competing paradigm gains strength and each of its premises becomes less vulnerable to revision. Coupled with this effect are the effects of new heuristic (2), where predictions made by the new paradigm are confirmed by revised premises. Each confirmed prediction can now add its support to the new paradigm, which in turn makes each supporting premise less vulnerable to revision. In short, our two induction heuristics reflect two directions in which paradigms can shift; heuristic (2) tends to build support for newer generalizations having little confirmed support but many predictions, while heuristic (1) tends to retain support for older generalizations having firm support. In our future research, we plan to use this general approach to model historical shifts in scientific paradigms, in particular the shift from phlogiston theory to oxygen theory.

Acknowledgments

This research was supported by Contract N00014–84–K–0345 from the Information Sciences Division, Office of Naval Research. We would like to thank Randy Jones, Bernd Nordhausen, and Paul O'Rorke for discussions that helped us form the ideas presented in this paper.

References

de Kleer, J. (1984). Choices without backtracking. *Proceedings of the Fourth National Conference on Artificial Intelligence* (pp. 79–85). Austin, TX: Morgan Kaufmann.

Doyle, J. (1979). A truth maintenance system. *Artificial Intelligence, 12*, 231–272.

Kuhn, T. S., (1970). *The structure of scientific revolutions.* Chicago, IL: University of Chicago Press.

Rose, D., & Langley, P. (1986a). STAHLp: Belief revision in scientific discovery. *Proceedings of the Fifth National Conference on Artificial Intelligence* (pp. 528–532). Philadelphia, PA: Morgan Kaufmann.

Rose, D., & Langley, P. (1986b). Chemical discovery as belief revision. *Machine Learning, 1*, 423–451.

Zytkow, J. M., & Simon, H. A. (1986). A theory of historical discovery: The construction of componential models. *Machine Learning, 1*, 107–136.