Scientific discovery: induction and beyond

Vincent Corruble

LIP6, Université Paris VI
4, Place Jussieu - Boîte 169
75252 PARIS Cedex 05
FRANCE
email: Vincent.Corruble@lip6.fr
Telephone: 33-1-4427-8800 Fax: 33-1-4427-7000

Abstract : This paper discusses the role of induction in medical discoveries and insists on the steps which are problematic from the inductive standpoint, in particular those steps concerning the construction of data and the evaluation of hypotheses. Building on earlier experimental work, we investigate here how alternative frameworks can help provide a better account of how distinct theories were proposed and why they were historically rejected before they got accepted. We conclude by outlining the limitations of the inductive framework *per se* as a single general framework accounting for human scientific discoveries.

Keywords : scientific discovery, induction, medicine

1. Introduction

A large part of the work done in Artificial Intelligence on Machine Discovery has followed a data-driven approach. How can a machine, given data similar to the one available to human scientists, produce a valuable scientific discovery ? This question naturally leads one to think in inductive terms, going from the data to a theory which has a wide applicability. Induction is the type of reasoning which has been the most widely studied by the machine learning community, and it is therefore a legitimate goal to estimate the significance of the intersection between the fields of Machine Learning and of Philosophy of Science. This question was considered in earlier work [Corruble & Ganasica 1996, Corruble & Ganascia 1997] that focused on the field of medicine. In this paper, we summarize this work which showed that, on two cases of important medical discoveries, an inductive framework can help the reconstruction of a large part of the reasoning process involved in history. More specifically, this reconstruction managed to account both for the reasoning that led to the actual discovery (ie the acceptation of the correct theory) and also, to a certain point, for the (erroneous) one that led earlier researchers to believe in alternate theories. We then go back to this earlier work and outline the limitation of the inductive framework for our task. These limitations are related to the fact that the inductive framework has a high potential to reconstruct the stage of hypothesis formation, but provide little insight into the stages of data construction on one hand, and of pursuit and evaluation of hypotheses on the other hand. We conclude by evaluating the potential of a mixed inductive/abductive framework to palliate the limitation of our inductive simulations.

2. A first simplistic view of induction in discovery

A first view of induction often displayed within machine learning is that theories appear by applying some form of inductive reasoning (such as supervised learning) onto data describing the phenomenon being studied. This view is often oversimplified but can lead to succesful application, both in terms of new discoveries and for the simulation of former discoveries. We used this first model (Figure 1) for the reconstruction of the discovery of scurvy (Corruble & Ganascia 1997). Based on 19th century data describing cases of scurvy, a general rule induction system managed to come up with the theory that was finally accepted in the 20th century, a theory blaming the lack of fresh fruits and vegetables as the only cause of the disease (Figure 2).

Figure 1: A first model of discovery based on induction

DIET (food *variety*):

It was J.F. Bachström (1734) who first expressed the opinion that, "Abstinence of vegetables is the only, the true, the first cause of scurvy." [**Erreur ! Source du renvoi introuvable.**MAHE, 1880]

Set I: Rules 3,4,8 use in their premises the variety of the diet.

```
R3:  IF diet-variety                 THEN disease-severity          [5]
R4:  IF diet-variety                 THEN disease-severity          [4]
R8:  IF diet-variety                 THEN disease-severity          [11]
```

Set II: Rules 7,10 use in their premises the presence (or absence) of fresh fruits and vegetables in the diet.

```
R7:  IF fresh_fruits/vegetables = no   THEN disease-severity        [5]
R10  IF fresh_fruits/vegetables = yes  THEN disease-severity        [13]
```

Figure 2: Excerpts from the 1880 medical encyclopedia
followed by the corresponding rules proposed by CHARADE.

However, this simulation did not account well for the intense debate that took place in the medical community between the 16th and 19th Century where the favored theory blamed humidity rather than diet. A good simulation should be able to account both for successful inference (the one that leads to the correct theory) and for erroneous reasoning. In a second experiment, we managed to simulate this reasoning by introducing a priori knowledge in our model. In this new model (Figure 3), the data used for induction is not the "raw data" but

rather data which has been enriched via a medical knowledge base and some deductive inferences. An interesting point about this a priori knowledge in the case of scurvy is that it was implicit, and therefore unquestioned (it could be labeled as a medical dogma rather than purely scientific knowledge). As a result, it did not appear explicitly in the medical texts addressing scurvy in the 19th century. We only found reference to it by looking at recent work on the history of medicine.

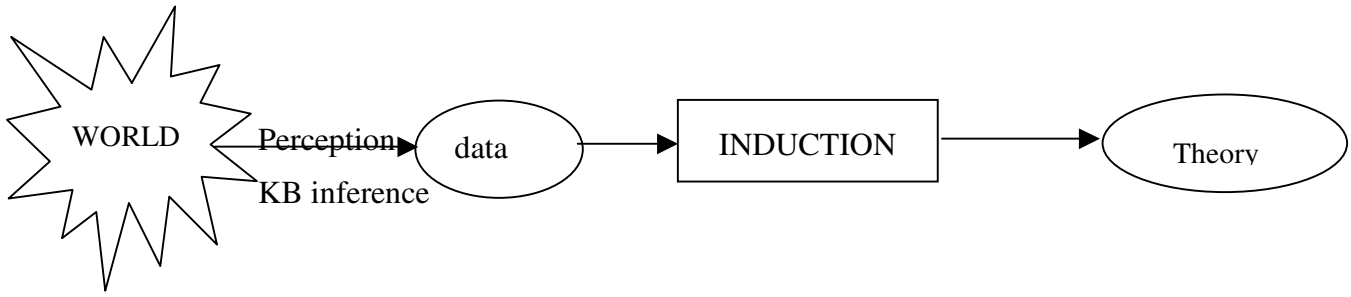

Figure 3: First enrichment of an induction-based model of discovery

A consequence of this work is that a model of discovery based on induction needs to be enriched to account for the complexity of human reasoning. The question that arises already is that of whether the whole process can be described as inductive. A positive answer is still acceptable because with scurvy, we are in a situation where the results of induction (here sets of rules) can be directly interpreted as potential theories on the cause of diseases. We see next a situation where the interpretation of the results of induction is itself a complex process which needs to be detailed.

3. Need for further enrichment of the discovery model.

We next conducted experiment [Corruble & Ganascia, 1996] on the discovery of the causes of leprosy. This discovery also has an intricate history, and the search for a better understanding of the disease is still an active research topic. To summarize, one can say that historically, two main theories were proposed: one that considered heredity as the main cause, one that considered contagion. In the 20th century, the infectious agent that causes the disease was isolated, but the key role of the immune system in the actual development of the disease has also appeared. Therefore current medical knowledge has links to both historical theories: the one blaming contagion, because of the infectious agent, and the one considering heredity, because it is believed that the characteristics of the immune system are at least partially inherited.

R1: IF father_affected = No THEN children = all_healthy
R2: IF { father_affected = No THEN children = some_sick
{ Mother_side = yes
{ age > 35
{ disease_type = aneasthetic
+ 3 minor rules

Figure 4. Model induced by PASTEUR

Based again on historical data from the 19th century, our simulation experiment aimed at reconstructing rationally the reasoning that could lead to correct or incorrect theories. We obtained interesting results (Figure 4) but yet met some difficulty in their interpretation. The main reason is that the data available described individual patients and gave only indirect information on how they contracted the disease (it was for example recorded if the parents, or the children, had the disease). This interpretation stage was performed "manually" but remains subject to criticism. To explain it more clearly requires a slightly more complex model (See Figure 5) of discovery which adds a stage between the results of induction (which we now call patterns) and the theory. We can say for example, based on the induced patterns of Figure 4, that the data contains evidence supporting the theory of heredity. Whether or not this evidence leads to acceptance of the theory is a complex problem which has been studied for instance in (Thagard,). The main point here is that the bridge between Patterns and Theory can only be crossed meaningfully with some background knowledge on the problem. In the next section we describe two approaches to tackle this problem. The first one led to a system while the 2nd is more of a research direction.

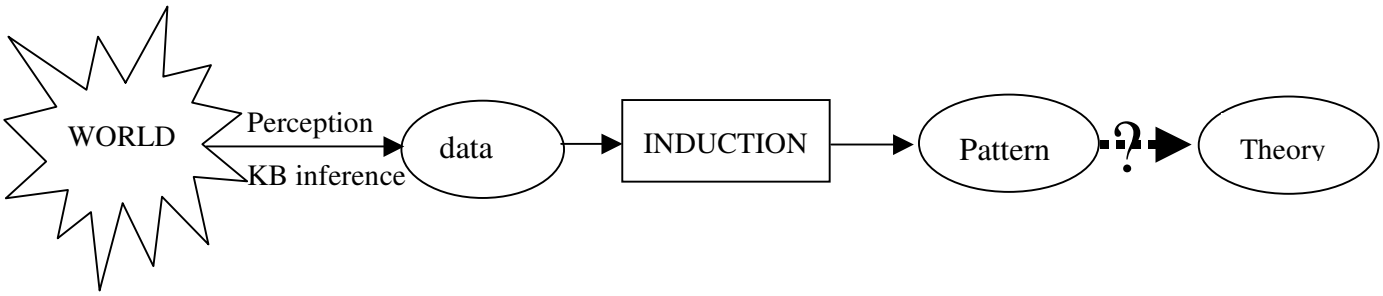

Figure 5: An induction(?)-based model enriched for data construction and results interpretation

4. Knowledge-based processing of induced patterns

The last stage of the complex process described in Figure 5 can be understood in the context of pre-existing background knowledge. We can in that context be interested in the interaction between this background knowledge and the induced patterns.

A first way to see this interaction is to consider that the pre-existing knowledge is dominant: in that context, induced patterns that are not consistent with this knowledge are subject to rejection.. We followed this approach in the design of the FILTER system (Corruble & Sleeman, 1998, Corruble 1999). This system can have applications in scientific discovery (we applied it to current medical research on severe head injuries) or more generally to the field of knowledge discovery in databases. Its basic assumption is that the human user cannot easily express the pre-existing knowledge used in the analysis of patterns, but can at least apply that knowledge to assign induced patterns to one of 3 categories: accepted (confirms or support current views), rejected (inconsistent with existing knowledge), innovative (neither inconsistent nor confirmed, i.e. a potential discovery…). The fact that a domain expert would reject a pattern induced from data could seem dubious, yet it is quite a common occurrence if one considers that the model of Figure 5 describes a number of stages that all involve a

number of subjective choices. Even the inductive step proper contains essential biases as is well-known in the field of Machine Learning. Once the patterns are classified in 3 categories, it is easier to diagnose how to adapt the earlier elements of the inductive chain, or to question the pre-existing knowledge used in this categorization.

On the other hand, the pre-existing knowledge can be suggestive instead of being dominant. In that case, this knowledge does not act as a filter anymore: the interaction is by essence a process of selection among the pre-existing theories, a process guided by the patterns inferred at the inductive stage. Though the filtering approach has proved useful in some applications, the suggestive view point is much more adapted to the problem posed by our work on leprosy. The theories of contagion, heredity, and later of immunity, are suggested by the pre-existing knowledge. The rules induced with the PASTEUR algorithm (Corruble et. al., 1995) provide support for one or more of these theories. However, this "suggestive" approach, to be useful, requires two stages which can cause problems. First this suggestive knowledge needs to be made explicit to start with. This should be relatively easy for a historical reconstruction (such as ours on leprosy) but can be very hard, or practically imporssible, on an ongoing research. It requires an unusual ability to stand back from one's own research. Second, it would be particularly useful to be able to formalize before the inductive step is carried out how various patterns would either reinforce or degrade our confidence in some of the pre-existing hypotheses.

5. Conclusion

Based on earlier experimental work, we have shown in this paper that the inductive framework, and tools provided by machine learning research, have a high potential for simulating important part of a data-driven discovery process. However we have also shown that for at least two key stages before and after the inductive step per se, a stage of data construction, and the stage of result interpretation, pure induction needs to be supplemented. Recent work in Machine Learning have addressed some of these limitations, and we have shown how we integrated new modules in our own simulations.

Though we have placed ourselves in an inductive framework in this paper, the reader probably realizes that our latest points concerning results interpretation fit particularly well within an abductive framework. The main difference with classical abduction here is that the abductive inference does not select a pre-existing hypothesis directly based on observational data, but rather based on induced patterns. One can note though that these patterns can also be seen as a way to summarize some complex observational data.

References

V. Corruble, Frédéric Thiré, J.G. Ganascia, Pasteur: a system for the induction of comprehensible hypotheses, AAAI-95, Workshop on Machine Learning and Comprehensibility. Montreal, 1995

V. Corruble and J.G. Ganascia, Discovery of the Causes of Leprosy: a Computational Simulation, in: Proceedings of 13th National Conference on Artificial Intelligence, Portland (OR), USA (1996) 731-736.

Vincent Corruble, Jean-Gabriel Ganascia, Induction and the discovery of the causes of scurvy: a computational reconstruction, Artificial Intelligence (91)2 (1997) pp. 205-223

V. Corruble, D. Sleeman: A knowledgeable system for knowledge discovery. Machine Intelligence 16, York 1998. Donald Michie and Stephen Muggleton

V. Corruble. Collaborative knowledge discovery with Pasteur and Filter: A case of mixed-initiative intelligence. AAAI-99 Workshop on Mixed-Initiative Intelligence

P. Thagard, Ulcers and bacteria I: Discovery and acceptance., Studies in History and Philosophy of Science (forthcoming).