# Abduction and Induction in Learning Task: which needs the other ?

**Engelbert Mephu Nguifo**[1]

**Abstract.** Abduction and induction are two basic and logical inference forms of reasoning. This paper addresses the intersting question of which form of reasoning needs the other in the context of learning task. It appears that if it is easy to show that abduction needs induction, it is also possible to prove that induction needs abduction in some sense.

## 1 INTRODUCTION

Recent years have witnessed a growing interest in developping multistrategy learning systems that integrate two or more inferences types and/or computational paradigms (ML, 1993). Such systems take advantage of the complementarity of different inference types of representational mechanisms. Michalski (1993) describes different types of inference encounter in machine learning (ML) systems. Among those inference types, there are induction and abduction which are, with deduction, the basic logical inferences. While deduction is tracing forward implications, abduction and induction are tracing backward implications, in such a way they could be viewed as some kind of 'reversed' deduction.

Induction is used to synthetize information from a set of observations into a general hypothesis. Abduction allows to assume a fact if it is a possible explanation for another fact to be explained.

Many works have been done in order to define both inferences (Pierce, 1898; Michalski, 1993; Ganascia, 1994; Torasso & al., 1995) and their usability in different domains where hypothetical reasoning is important (AI, 1990; Barboux, 1990; Codognet, 1994; Satoh, 1994; Eiter, 1995). This paper focusses on a somehow different point which consists in finding relationships between abduction and induction. Console & al. (1991) describes similar work, however with abduction and deduction; They showed that abduction can be reduced to deduction on a transformed domain theory.

This paper deals with the question setted up in the title, which is about the necessity of using abduction (respectively induction) before proceed with induction (resp. abduction) in the domain of learning tasks. Hypothetical reasoning is important in such domain, so that different systems arisen from machine learning, integrate both basic inferences.

Michalski (1993) defines the task of learning as respectively inferencing and memorizing. If this could be seen as a basic definition of learning, I believe that the ML systems should not be reduced to derive new knowledge from input knowledge and memorizing it, but also they should be able to explain derived knowledge, and to interact with an expert-human or other AI systems.

This paper is divided in 4 sections. The section 2 describes what is learning task. Then the section 3 defines abduction and induction. Section 4 shows that abduction needs induction, and section 5 describes how induction needs abduction.

## 2 WHAT IS LEARNING TASK ?

We have proposed a modelling framework that integrate similarity-based learning (SBL) and explanation-based learning (EBL) approaches (Mephu, 1995). An implementation of this modelling framework was done towards the LEGAL system (Mephu, 1994). This modelling framework is based on the fact that the whole learning process is a successive application of basic inferences mechanisms, which are abduction, induction, deduction, with validation and revision.

In practice, to gain confidence of a human-expert, a ML system should always provided results expected by the expert. However the expert beliefs may be wrong as it is sometimes the case in domain with imperfect theory. In such case, the argumentation of system results could allow the expert to change its belief possibly after some experimentation. For example, when dealing with imperfect domain theory, initial data may be incomplete or may contain many biases. Knowing this, the expert can modify data that he provides to the system in order to allow it to refine its built knowledge. Thus the goal of ML systems should be not only to find a formulation of the concept, but also to help the user to discover data mismatches. Such goal explains partly the growing interest in multistrategy learning systems.

The learning task (Mephu, 1995) is achieved by dividing it in five subtasks. Considering an initial goal given by the human-expert, the *first* subtask is to search for a context (a set of examples and their description) which could describe the concept associated to the goal. The first subtask is achieved by choosing the *examples* characterizing a concept, and in finding the language used to describe examples. Examples are described by means of *facts*. The *second* subtask consists in finding a generalisation of the context. This generalisation is then used by the *third* subtask to produce new result as prediction and argumentation of new data. A comparison of the result obtained and the initial goal is done by the expert in the *fourth* subtask, and could generate the *fifth* subtask which consists in refining knowledge provided by the three first subtasks.

Let us illustrate the learning process through a little concept example: "To be a good General". For example, the concept "a good general" can be associated to the goal "How to characterize a good general in the army". Considering this example, the expert gives different statements which could help the system to find the initial context. For example, "I know that Napoleon is a good general. He is small, clever, fat, white. He has one arm, two legs. He has done war. He has won war. ..."; "I know that Wellington is

[1] CRIL - IUT de Lens - Université d'Artois, Rue de l'université SP 16, 62307 Lens cedex, France. mephu@lifl.lifl.fr

a good general. He is young, fat, white. He has two arms, one leg. He has done war. He has won war. ...";. By examining these kinds of statements, the system infers (by abduction) that the context may contain Napoleon, Wellington and so on as examples described by the binary facts: small, clever, has-won-war, fat, white ... An inductive generalization of this context can be: "white and. has-won-war". This characterization could then be sucessfully used to recognize new good generals until it fails when looking at 'Soundiata Keita' who wasn't white, had won war and was considered in the last century as a good general. An explanation of this failure comes from the fact 'white' in the characterization. Knowing this, the system can modify its background knowledge to the simply fact 'has-won-war'. This refinement allows concept formation context as developped in (Wrobel, 1994).

Therefore, learning should inevitably integrate other tasks as explanation, interaction, and so on. From this point of vue, could we say that induction needs abduction, or on the contrary that abduction needs induction? To answer this question, I need the definitions of both abduction and induction.

## 3 DEFINITION OF ABDUCTION AND INDUCTION

This section gives a general definition of what are abduction and induction. A small example is taken to illustration those definitions.

There are different kinds of implementation of abduction and induction in the literature (AI, 1990; Barboux, 1990; Codognet, 1994; Satoh, 1994; Eiter, 1995). According to the philosophical definition of Peirce's (1898) syllogistic classification, we can illustrate these two inferences through a small example as follow:

- Knowing the two facts:
  - (a) "*All Brazilian are passionately fond of soccer*" and
  - (b) "*Paul is passionately fond of soccer*",

  we infer, by **abduction**, the fact:
  - (c) "*Paul is Brazilian*".

- Conversely knowing that:
  - (d) "*Paul is Brazilian*", and
  - (e) "*Paul is passionately fond of soccer*",

  we infer, by **induction**, that
  - (f) "*All Brazilian are passionately fond of soccer*".

Abduction often refers to explanation (Torasso, 1995) (the fact (c) explains the fact (b) with fact (a) in the background knowledge). While induction generally refers to generalisation (going through one entity - fact (d) - which verifies one property - fact (e) -, we generalize by the fact (f)).

Both two are inferences that incorporate a certain form of "reversed" deduction. This rejoints Michalski's (1993) point of vue for which abduction is based on tracing backward domain-dependent implications, while induction is based on tracing backward a domain-independent implication.

## 4 ABDUCTION NEEDS INDUCTION ?

According to the previous definition, abduction can be formally illustrated as follows:

| | |
|---|---|
| Input knowledge (I): | q(a) |
| Background knowledge (BK): | p(x)->q(x) |
| ---------- | |
| Output knowledge (O): | p(a) |

where p(x) and q(x) are two properties which, in the previous example, respectively mean "x belongs to the set X - the set of brazilian", and "x is passionately fond of soccer".

How do we obtain the BK: p(x) -> q(x)? Two cases may be considered. The first one is that the BK comes from inferences done by the ML system. In such case, BK is derived from many observations as p(b), p(c), q(c), q(b). Thus the ML system cannot explain, using abduction, without using inductive inference to generate BK.

The second case is when BK comes from outside, i.e. BK is given by the user of ML system which may be the case with EBL systems. Nevertheless, the user build himself the BK by induction from observations he made.

So induction is implicit in the latter case, as it is manually done by the user, while it is explicit in the former one as done by the ML system. Therefore, as we need BK to proceed abduction, we cannot make any abduction without the implicit or explicite use of induction.

## 5 INDUCTION NEEDS ABDUCTION ?

Let us consider the case of induction here. Induction can be formally illustrated as follows:

| | |
|---|---|
| Input knowledge (I): | q(a) |
| Background knowledge (BK): | p(a) |
| ---------- | |
| Output knowledge (O): | p(x)->q(x) |

How do we obtain the input knowledge and BK as data for inductive inference. More precisely, how do we initiate the initial context for inductive inference? The initial context can be setted up by the expert or another ML system from abductive inference as it is shown in section 2. Considering also the example of section 3, the expert uses different kinds of statements to find the initial context. For example, "I know that Paul is brazilian. He is passionately fond of soccer. ..."; "I know that John is brazilian. He is passionately fond of soccer. ...";. By examining these kinds of statements, the expert infers, by abduction, that the context may contain Paul, John and so on as examples described by the binary facts: to be brazilian, and to be passionately fond of soccer ...

Here, input and BK are observations used to derived the general rule: p(x) -> q(x). When this knowledge is used to derived new results (prediction, diagnosis, ...), the ML system should be able to explain the results provided. The explanation is done by using abductive inference. For example, when predicting that a new person, John, who is brazilian, is passionately fond of soccer by using learned knowledge p(x)->q(x), the ML system should be able to derived that John is a Brazilian. This comes out from our definition of the learning process which is enlarged to explanation task.

We consider that finding an initial context (for example, a set of brazilians and their descriptions) which corresponds to a goal (for example, characterizing a brazilian) can be termed abduction. The fact that we are searching for a characterization of what is a brazilian, is explaining by given a set of observations made from people of Brazil. Therefore, we cannot make any induction without the implicit or explicit use of abduction.

## 6 CONCLUSION

The goal of this paper has been to advocate that in the ML systems, abduction can not be used without induction, and induction needs abduction. Now as these links are exhibited, if abduction (resp. induction) needs induction (resp. abduction) to operate, then which of the two should we firstly use in learning process?

It comes from this question that interaction with an expert or another system is necessarily to set up initial data, depending on the problem which is to be addressed.

This question can bring out another more familiar and intractable question which is:

*'which one of man or woman was the first in the earth'*

## ACKNOWLEDGEMENTS

Financial support for this work has been partially made available by the Ganymede project of the Contrat de Plan Etat/Nord-Pas-de-Calais.

## REFERENCES

AI. 1989. *Artificial Intelligence*, Special vol. on Machine Learning. Ed. Carbonnell, J.G., 40(1)-3.

Barboux, C., and Sallantin, J. 1990. Contrôle par objection d'une théorie semi-empirique. *Revue d'Intelligence Artificielle*, 4(2):29-41.

Codognet, C., and Codognet, P. 1994. Abduction and Concurrent Logic Languages. In Proceedings of 11th ECAI, Amsterdam (The Netherlands), 75-79, John Wiley & Sons, Ltd.

Console, L., Theseider Dupré, D. and Torasso, P. 1991. On the relationship between abduction and deduction. *Journal of Logic Computation*, 1(5):661-690.

Ganascia, J.G. 1994. Logique et Induction: un vieux debat. Laforia Internal report. University of Paris VI.

Hayes-Roth, F. 1986. Using Proofs and Refutations to Learn from Experience. *Machine Learning : An A.I. Approach*, T.M.Mitchell & al. Eds, San Mateo, CA, p221-240.

Lakatos, I. 1984. *Preuves et Refutations*. Hermann Ed.

Mitchell, T. 1982. Generalization as Search. *Artificial Intelligence*, 18(2):203-226.

ML, 1993. *Machine Learning: Special Issue on Multistrategy Learning,* 11(2-3). Michalski, R.S. guest Editor.

Mephu Nguifo, E. 1994. Galois Lattice: a Framework for Concept Learning. Design, Evaluation and Refinement. In Proceedings of 6th Intl. Conf. on Tools with Artificial Intelligence, New Orleans (LA), November 6-9 th, IEEE Press.

Mephu Nguifo, E. 1995. Multi-Level Concept Modelling for SBL. In Proceedings of MLNet (ECML) workshop. on Knowledge Level and Machine Learning, Heraklion (Crete - Greece), April 28-29th, Eds. Dieter Fensel & al.

Michalski, R. S. 1993. Inferential Theory of Learning as a Conceptual Basis for Multistrategy Learning. *Machine Learning* 11(2-3): 111-151.

Newell, A. 1982. The knowledge level. *Artificial Intelligence*, 18:87-127.

Peirce, C.S. 1898. *Reasoning and the Logic of Things*, Cambridge, Massachusetts: Havard University Press.

Quinlan, J.R. 1986. Induction of Decisions Trees. *Machine Learning*, 1:81-106. T.M.Mitchell & al. eds.

Russell, B. 1956. *The Principles of Mathematics*. Allen Unwin Ed., London.

Satoh, K. 1994. A Top Down Proof Procedure for Default Logic by Using Abduction. In Proceedings of 11th ECAI, Amsterdam (The Netherlands), 65-69, John Wiley & Sons, Ltd.

Torasso, P., & al., 1995. On the role of Abduction. *ACM Computing Surveys*, 27(3):353-355.

Wrobel, S. 1994. *Concept Formation and Knowledge Revision*. Hermann Ed.