Abduction and Induction: Syllogistic and Inferential Perspectives

Peter A. Flach[1]

Abstract. Abduction and induction are reasoning forms for drawing conclusions from incomplete information. Induction, i.e. inferring properties of sets of individuals from properties of individuals, was already distinguished by Aristotle, while the term 'abduction' was introduced much later by Peirce for inference of explanations for observed phenomena. Both reasoning forms are presently being studied and applied by researchers in artificial intelligence and logic programming. However, the current views of abduction and induction and their interrelation are problematic, which is mainly caused by the fact that Peirce developed **two** perspectives on abduction, one based on syllogisms, the other on the underlying inferential pattern. In this paper I argue that both perspectives have their merits but need further formalisation. Furthermore, I propose a formalisation of the inferential perspective.

1 INTRODUCTION

Abduction and induction are key processes in intelligent reasoning, enabling an agent to deal with incomplete information. Induction is the oldest and most controversial of the two, provoking heated debate about such fundamental questions as: How are inductive conclusions justified? Is there a logic of induction? Does induction form the basis for scientific hypothesis formation? In fact, there does not even seem to be a generally accepted **definition** of induction. For the moment, let us adopt the following (which is only a slightly more precise paraphrase of the popular slogan 'induction is inference from the particular to the general') as one of the characteristic properties of inductive arguments: the premisses are statements about certain individuals, and the conclusion is a statement about sets of such individuals. Clearly, this characteristic stands in need of a specification of the exact relation between premisses and conclusion, which is one of the goals of this paper.

Abduction is a form of reasoning commonly described as 'inference to the best explanation'. It was introduced less than a century ago by the American philosopher and logician Charles Sanders Peirce. In the last 15 years abduction has become quite popular in logic programming, mainly because as a generic form of inference it encompasses other forms like reasoning by default and negation as failure, and can be used as the basis for logic programming semantics [1].

However, once one steps outside of the context of logic programming there are again terminological problems, not only concerning abduction but also its relation to induction. For instance, Josephson writes that 'it is possible to treat every good (…) inductive generalization as an instance of abduction' [2, p.19], while Michalski has it that 'inductive inference was defined as a process of generating descriptions that imply original facts in the context of background knowledge. Such a general definition includes inductive generalization and abduction as special cases' [3, p.188].

The conflicting views of these two authors can be traced back to the **two** theories of abduction developed by Peirce. Since it is not widely known that Peirce came to believe that his early theory was flawed, I will devote section 2 to an exposition of the two perspectives: the *syllogistic* theory and the *inferential* theory (my terminology). The syllogistic perspective is employed in logic programming contexts, and will be briefly discussed in section 3, where I argue that the syllogistic perspective is intuitively justifiable, but that these intuitions are not yet adequately reflected in the semantics employed. The inferential viewpoint has to give up the distinction between abduction and induction as separate forms of reasoning, but allows for an attractive formalisation of the inference patterns underlying hypothetical reasoning in terms of *conjectural consequence relations* (section 4). This formalisation is a first step towards a 'logic of discovery', strongly related to recent theories about nonmonotonic reasoning [4]. Finally, I will discuss the results of the foregoing analysis in section 5.

2 PEIRCE'S THEORIES OF ABDUCTION

Peirce was a very prolific thinker and writer, but virtually nothing was published during his life. His collected works [5] therefore reflect, first and foremost, the evolution of his thinking, and should be approached with some care. With respect to abduction and induction Peirce went through a substantial change of mind during the decade 1890 – 1900 [6]. Below I will briefly explicate both his early, syllogistic theory, on which the current view of abduction in logic programming is based, and his later, inferential theory, which treats abduction as the process of hypothesis formation.

2.1 The syllogistic theory

Peirce's first classification of reasoning forms follows Aristotle in employing syllogisms. Peirce uses the following example, embodying a syllogism known as *Barbara*:

All the beans from this bag are white.
These beans are from this bag.
∴ These beans are white.

While this syllogism is deductively valid, two other non-deductive syllogisms are obtained from Barbara if we exchange the conclusion (or *Result*, as Peirce calls it) with either the major premiss (the *Rule*) or the minor premiss (the *Case*):

Case. — These beans are from this bag.
Result. — These beans are white.
∴ *Rule*. — All the beans from this bags are white.

Rule. — All the beans from this bag are white.
Result. — These beans are white.
∴ *Case*. — These beans are from this bag.

The first of these two syllogisms (inference of the *rule* from the *case* and the *result*) is readily identified as constituting an inductive argument, generalising from specific beans to all the beans in the bag. The sort of inference exemplified by the second syllogism (inference of the *case* from the *rule* and the *result*) Peirce calls *making a hypothesis* or, briefly, hypothesis [5, 2.623] — the term 'abduction' is introduced only in his later theory.

[1] INFOLAB, Tilburg University, POBox 90153, 5000 LE Tilburg, the Netherlands. Email: Peter.Flach@kub.nl.

© 1996 P. Flach
ECAI 96. workshop on Abductive & Inductive Reasoning

2.2 The inferential theory

Around 1900 Peirce discovered that he was 'too much taken up in considering syllogistic forms' [5, 2.102]. Instead, he started to perceive the three stages of scientific inquiry as three distinct forms of reasoning. Peirce calls the first stage, coming up with a hypothesis to explain the initial observations, *abduction*; predictions from a suggested hypothesis are derived by *deduction*; and estimating the credibility of that hypothesis through its predictions is *induction*.

2.2.1 Abduction

According to Peirce's later theory, abduction is the process of forming an explanatory hypothesis, given a number of observations. This process is not algorithmic: 'the abductive suggestion comes to us like a flash. It is an act of *insight*, although of extremely fallible insight' [5, 5.181]. On the other hand, abduction

> 'is logical inference (…) having a perfectly definite logical form. (…) Namely, the hypothesis cannot be admitted, even as a hypothesis, unless it be supposed that it would account for the facts or some of them. The form of inference, therefore, is this:
>
> The surprising fact, C, is observed;
> But if A were true, C would be a matter of course,
> Hence, there is reason to suspect that A is true.'
> [5, 5.188–9]

Peirce thus clearly makes a distinction between the declarative view and the procedural view of reasoning, anticipating the modern distinction between semantics and proof theory. Abduction has a well-defined semantics, even if humans don't actually employ a proof procedure implementing that semantics.

Many authors have translated the logical form of abduction given by Peirce into the following inference rule:

$$\frac{C \ , \ A \models C}{A} \qquad (1)$$

However, clearly this inference rule is no more a **definition** of abduction than *Modus Ponens* is a definition of deduction. In section 4 I will demonstrate how a system of inference rules, similar to this one but employing the notion of a conjectural consequence relation, does establish a definition of abduction.

To sum up, in his inferential theory Peirce defines abduction as the process of forming a hypothesis that explains given observations. Contrary to the syllogistic perspective, no restrictions are put on the form of the hypothesis: it may be either a particular fact or a general rule. Furthermore, by taking logical implication for explanation Peirce identifies abduction with reversed deduction. In section 4 I will demonstrate how this can be generalised to other notions of explanation.

2.2.2 Induction

Induction is identified by Peirce as the process of confronting a hypothesis, through selected predictions, with reality:

> 'When I say that by inductive reasoning I mean a course of experimental investigation, I do not understand experiment in the narrow sense of an operation by which one varies the conditions of a phenomenon almost as one pleases. (…) An experiment (…) is a question put to nature. (…) The question is, Will this be the result? If Nature replies "No!" the experimenter has gained an important piece of knowledge. If Nature says "Yes," the experimenter's ideas remain just as they were, only somewhat more deeply engrained.' [5, 5.168]

This perception of induction as estimating the extent to which a hypothesis is confirmed by evidence was later developed by Carnap into a system of 'inductive logic' [7]. One should note that induction, thus perceived, is not **reasoning** (in the sense of inference of a conclusion from premisses), since the hypothesis is already inferred abductively. In Peirce's words: 'Induction consists in starting from a theory, deducing from it predictions of phenomena, and observing those phenomena in order to see *how nearly* they agree with the theory'; it 'does nothing but determine a value' [5, 5.170–1].

In summary, in Peirce's later theory the logical part of scientific reasoning is entirely carried out by abduction (to come up with a hypothesis) and deduction (to make predictions), and the non-logical part is left to induction (to determine a confirmation measure). In section 4 I will demonstrate how a qualitative approach to confirmation can lead to an alternative logical account of hypothesis formation.

3 THE SYLLOGISTIC PERSPECTIVE IN LOGIC PROGRAMMING

Although Peirce dismissed his earlier syllogistic perspective on abduction and induction, it is widely applied in logic programming. This is understandable, since definite clause logic has such a strong syllogistic flavour: a rule corresponds to a non-ground definite clause with non-empty body, a case corresponds to a ground fact, and a result corresponds to a ground query. Peirce's abductive syllogism then amounts to inference of a ground fact that enables a ground query to be proved from a given definite clause, while the inductive syllogism amounts to inference of a definite clause that enables a ground query to be proved from a given ground fact.

However, clearly definite clause logic is much more powerful than syllogistic logic, including non-monadic predicates, function symbols, and sets of (possibly recursive) rules. Because of this extra power the neat syllogistic distinction between abduction and induction breaks down. For instance, consider the following clauses:

```
flies(X):-bird(X),flyingbird(X).
bird(polly).
```

The first clause represents the default rule 'birds typically fly'; the abducible predicate `flyingbird` is used to name instances of this rule [8]. The observation `flies(polly)` is thus abductively explained by `flyingbird(polly)`, indicating an application of the default rule. But for a fixed language the set of possible clauses is denumerable, so the same trick may be used to name every possible inductive hypothesis, effectively emulating induction by abduction.

Conversely, in inductive program synthesis the intended program is typically recursive and thus may include a ground base case such as `reverse([],[])`. When extensional coverage is employed (as in FOIL [9]) it is quite possible that the recursive clause has already been inferred, so that inference of the base case from a ground query and a recursive clause is effectively an abductive step.

Even though the distinction between abduction and induction in logic programming is more contrived than Peirce's syllogisms suggest, many people feel that the intuition underlying the distinction is quite straightforward. In abduction we try to complete our knowledge about a certain individual, while in induction we have complete knowledge about certain individuals but want to extend this knowledge to other individuals. An inductive hypothesis

is predictive in the sense that it equally applies to individuals unknown at the time the hypothesis was put forward; an abductive hypothesis is only predictive with regard to the individual about which the observation was made.

Straightforward as the intuition may be, it has as yet not been adequately reflected in an appropriate semantics. Clearly neither Tarskian semantics nor Herbrand models can accomodate the two kinds of predictiveness hypotheses may display. Yet it is at the level of semantics, rather than at the level of syllogisms, that the distinction between abduction and induction must be made. We consider the development of such a semantics for abduction and induction as an important open problem.

4 THE INFERENTIAL PERSPECTIVE: A FORMALISATION

We now make a switch from the syllogistic to the inferential perspective. The main benefit of Peirce's inferential perspective is that it makes a clear distinction between the logical process of hypothesis formation and the extra-logical process of hypothesis evaluation. In this section I will present an axiomatic account of the logic of hypothesis formation. The approach has been inspired by the seminal work of Kraus, Lehmann & Magidor on the formalisation of nonmonotonic reasoning [4]. The main concept in their framework is the notion of a consequence relation $\vdash$, which they treat as a binary metalevel predicate. Analogously, I will employ the notion of a *conjectural consequence relation*, denoted by $\mid\!\sim$: a metalevel statement $\alpha \mid\!\sim \beta$ means that β is a possible conjecture (or hypothesis) given evidence α. Furthermore, I will employ the monadic metalevel predicate $\models$, indicating validity as usual (i.e. truth in all models; an implicit background theory may be taken into account by restricting the set of models). Metavariables α, β, γ, δ, … range over formulæ of the object-level language *L*. In general *L* stands for an arbitrary first-order predicate language; however, the formal results below only hold for propositional languages.

4.1 Explanatory reasoning

Following Peirce we might take the following as a basic property of conjectural consequence relations:

Converse Entailment: $$\frac{\models \beta \rightarrow \alpha}{\alpha \mid\!\sim \beta}$$

This property 'lifts' inference rule (1) to the metalevel, stipulating that it is sufficient for explanatory hypotheses to entail the evidence. Further obvious properties include the principles of verification and falsification, well-known from philosophy of science:

Verification: $$\frac{\models \alpha \wedge \beta \rightarrow \gamma \ , \ \alpha \mid\!\sim \beta}{\alpha \wedge \gamma \mid\!\sim \beta}$$

Falsification: $$\frac{\models \alpha \wedge \beta \rightarrow \gamma \ , \ \alpha \mid\!\sim \beta}{\alpha \wedge \neg\gamma \not\mid\!\sim \beta}$$

In these rules γ is a *prediction* drawn from conjecture β — if the prediction is indeed observed, the conjecture is still tenable, while if the prediction turns out to be false, the conjecture is refuted.

Falsification can be equivalently stated as follows:

Consistency: $$\frac{\alpha \mid\!\sim \beta}{\not\models \beta \rightarrow \neg\alpha}$$

The property of Consistency has a few technical consequences. For one thing, we have reflexivity only for consistent formulæ. Given a conjectural consequence relation $\mid\!\sim$, we call a formula α *admissible* iff $\alpha \mid\!\sim \alpha$. Whenever a property requires consistency of a formula this will be expressed by an admissibility condition of the form $\alpha \mid\!\sim \alpha$. For instance, Converse Entailment should be weakened in order to avoid inconsistent conjectures, as follows:

Admissible Converse Entailment: $$\frac{\models \beta \rightarrow \alpha \ , \ \beta \mid\!\sim \beta}{\alpha \mid\!\sim \beta}$$

As noted before, this property expresses that consistent conjectures entailing the evidence count as explanatory hypotheses. It can be derived[2] from the following stronger property:

Admissible Right Strengthening: $$\frac{\models \gamma \rightarrow \beta \ , \ \alpha \mid\!\sim \beta \ , \ \gamma \mid\!\sim \gamma}{\alpha \mid\!\sim \gamma}$$

This property states that a possible explanatory hypothesis β can be logically strengthened as long as it remains admissible (i.e. consistent), a well-known principle from machine learning.

The next two properties for explanatory reasoning establish a principle of compositionality regarding the evidence, i.e. $\alpha \wedge \beta \mid\!\sim \gamma$ iff $\alpha \mid\!\sim \gamma$ and $\beta \mid\!\sim \gamma$. The first property states one half of this principle, while the second takes a slightly more general form:

Additivity: $$\frac{\alpha \mid\!\sim \gamma \ , \ \beta \mid\!\sim \gamma}{\alpha \wedge \beta \mid\!\sim \gamma}$$

Incrementality: $$\frac{\models \alpha \rightarrow \beta \ , \ \alpha \mid\!\sim \gamma}{\beta \mid\!\sim \gamma}$$

Both properties are of considerable importance in machine learning, since they allow for an incremental approach. Additivity states that pieces of evidence can be dealt with in isolation. Another way to say the same thing is that the set of evidence explained by a given conjecture is conjunctively closed. Incrementality states a monotonicity property of induction, which can best be understood by considering its contrapositive: a conjecture that is rejected on the basis of evidence β cannot become feasible again when stronger evidence α is available. In other words: in incremental hypothesis formation the process of rejecting a conjecture is not defeasible (i.e. based on assumptions), but based on the evidence only. This is the analogue of the monotonicity property of deduction.

Regarding the semantic characterisation of explanatory consequence relations some results have been obtained [10]. The basic idea is to assume an *explanation mechanism* $\vdash$, and to define

$$\alpha \mid\!\sim \beta \text{ iff } Cn_{\vdash}(\alpha) \subseteq Cn_{\vdash}(\beta) \subset L \qquad (2)$$

where $Cn_{\vdash}(\alpha) = \{\gamma \mid \alpha \vdash \gamma\}$ is the *explanatory power* of α. In words, inference of conjecture β from evidence α should preserve explanatory power. The main question is then how properties of $\vdash$ affect properties of $\mid\!\sim$. A representation result has been obtained for $\models$ as explanation mechanism, in which case the consequence relation defined by (2) is called *strong explanatory*.

Theorem 1 *A consequence relation is strong explanatory iff it satisfies Verification, Falsification, Admissible Right Strengthening, Additivity, Incrementality, and the following two rules:*

Explanatory Reflexivity: $$\frac{\alpha \mid\!\sim \alpha \ , \ \neg\beta \not\mid\!\sim \alpha}{\beta \mid\!\sim \beta}$$

Conditionalisation $$\frac{\alpha \mid\!\sim \beta \wedge \gamma}{\beta \rightarrow \alpha \mid\!\sim \gamma}$$

Explanatory Reflexivity represents a weakened form of reflexivity, and is needed to ensure the connection between admissibility and

[2] The derivation requires Explanatory Reflexivity (see below).

consistency. Conditionalisation expresses a principle well-known from machine learning: if α represents the classification of an instance and β its description, then we may either induce a concept definition from examples of the form $\beta \rightarrow \alpha$, or we may add β to the background theory and induce from α alone. Since in our framework background knowledge is included implicitly, β is added to the hypothesis instead.

Representation theorems for weaker explanation mechanisms are still under study.[3]

4.2 Confirmatory reasoning

Not every useful hypothesis counts as an **explanation** of observed phenomena. For instance, the conjecture 'investments decrease when interests increase' does not explain the current investment rate given the interest rate, but hints at a certain relationship between the development of investment rate and interest rate. We might say that such conjectures are *confirmed* by the data.

Pioneering work on the formalisation of the relation of confirmation has been done by Carl G. Hempel, whose adequacy conditions for a material definition of confirmation can be expressed almost directly in the axiomatic framework developed here [11, pp.103–106, 110]:

(H1) any sentence which is entailed by an observation report is confirmed by it.

(H2) if an observation report confirms every one of a class K of sentences, then it also confirms any sentence which is a logical consequence of K.

(H3) every logically consistent observation report is logically compatible with the class of all the hypotheses which it confirms.

(H4) if an observation report B confirms a hypothesis H, then any observation report logically equivalent with B also confirms H.

(H1) expresses that logical entailment is a special case of confirmation. (H2) is quite reasonable if K contains a single hypothesis, but much less so in the general case since it basically states that the conjunction of all confirmed hypotheses is itself confirmed. (H3) suffers from the same problem: it disallows the possibility that two confirmed hypotheses partly contradict each other. (H4) finally expresses that the syntactical form of the evidence is immaterial.

Although Hempel's adequacy conditions are thus not unproblematic, they can be intuitively justified in certain cases. Before discussing this I will first reformulate (H1–4) in terms of conjectural consequence relations. To keep in line with the axiomatisation of explanatory reasoning I will treat inconsistent evidence as confirming no hypothesis at all, which requires a small change of (H1) and (H3). (H1) is translated as follows:

Admissible Entailment: $$\frac{\models \alpha \rightarrow \beta \ , \ \alpha \mathrel{|\!\sim} \alpha}{\alpha \mathrel{|\!\sim} \beta}$$

(H2) corresponds to the following two properties:

Right Weakening: $$\frac{\models \beta \rightarrow \gamma \ , \ \alpha \mathrel{|\!\sim} \beta}{\alpha \mathrel{|\!\sim} \gamma}$$

Right And: $$\frac{\alpha \mathrel{|\!\sim} \beta \ , \ \alpha \mathrel{|\!\sim} \gamma}{\alpha \mathrel{|\!\sim} \beta \wedge \gamma}$$

For (H3) it is sufficient to require consistency of evidence and single confirmed hypotheses, as expressed by the aforementioned rule of Consistency, since by Right And the conjunction of all confirmed hypotheses is itself confirmed. Finally, (H4) is easily expressed as

Left Logical Equivalence: $$\frac{\models \alpha \leftrightarrow \beta \ , \ \alpha \mathrel{|\!\sim} \gamma}{\beta \mathrel{|\!\sim} \gamma}$$

We thus see that the main problem with (H2) and (H3) can be traced to Right And, which is certainly too strong in the general case: the evidence may be simply too weak to decide between two incompatible hypotheses. In other words, Right And represents a *completeness assumption* with regard to the evidence, in the sense that α contains all the possible states of affairs that are somehow relevant for the domain. For instance, when talking about ravens and their colours such complete evidence would contain some black ravens, but also a few non-ravens, some of which are black and some of which are not — in such cases 'what you are not told looks like what you are told' ([12], p.149). This is reminiscent of the Closed-World Assumption, and indeed several authors have applied the CWA in order to form inductive hypotheses that are confirmed by the evidence, yet do not necessarily explain them [12, 13, 14].

As for the semantic characterisation of confirmatory reasoning, some initial ideas and results have been developed [10]. The main idea is to construct, from the evidence α, a set of *regular* interpretations $Reg(\alpha)$, which exhibit the same regularities as the evidence, and to define

$$\alpha \mathrel{|\!\sim} \beta \text{ iff } \varnothing \subset Reg(\alpha) \subseteq Mod(\beta) \qquad (3)$$

where $Mod(\beta)$ denotes the set of models of β. In words, the hypothesis should be true in every regular interpretation. Note that such regular interpretations are not necessarily models of the evidence. For instance, in the ravens example we may interchange the names of ravens and non-ravens without invalidating the inductive conclusion. This idea seems appealing because it allows a distinction between what is essential in the evidence and what is accidental.

However, in my initial investigations I have followed the more traditional route of assuming a preference ordering on models, and to define the most preferred models of the evidence as regular. In that case the consequence relation defined by (3) is called *preferential confirmatory*. This approach has been inspired by the aforementioned authors who employed a form of preferential reasoning by applying the CWA. The following representation result[4] is therefore a close variant of Kraus *et. al*'s characterisation of preferential reasoning [4].

Theorem 2 *A consequence relation is preferential confirmatory iff it satisfies Right Weakening, Right And, Consistency, Left Logical Equivalence, and the following three rules:*

Confirmatory Reflexivity: $$\frac{\alpha \mathrel{|\!\sim} \alpha \ , \ \alpha \not\mathrel{|\!\sim} \neg \beta}{\beta \mathrel{|\!\sim} \beta}$$

Cautious Monotonicity: $$\frac{\alpha \mathrel{|\!\sim} \beta \ , \ \alpha \mathrel{|\!\sim} \gamma}{\alpha \wedge \beta \mathrel{|\!\sim} \gamma}$$

Left Or: $$\frac{\alpha \mathrel{|\!\sim} \gamma \ , \ \beta \mathrel{|\!\sim} \gamma}{\alpha \vee \beta \mathrel{|\!\sim} \gamma}$$

Confirmatory Reflexivity represents a weakened form of reflexivity, similar to Explanatory Reflexivity. Cautious Monotonicity states that the assumptions embodied in two different conjectures do not

[3] As Daniel Lehmann observed, any inductive consequence relation defined by (2) will satisfy transitivity (Lehmann, personal communication).

[4] The semantic construction is a little bit more involved than suggested here; for details refer to [4] and [10].

contradict each other. This is a strengthening of the previously-mentioned rule of Verification, which only allows $\alpha \wedge \beta \mid\!\sim \gamma$ if β is conjectured on the basis of the same assumptions as γ. Left Or can be seen as a weakening of Incrementality, which clearly is invalid if we need to make completeness assumptions on the evidence.

The axiomatisation of the more general notion of regular interpretations is currently being studied. Note that Right And is a direct consequence of (3), which is therefore to be amended if we wish to get rid of Right And. One option is to require that at least one of the regular interpretations is a model of the hypothesis [10].

5 DISCUSSION AND CONCLUSIONS

The relation between the inferential and syllogistic perspectives may be summarised as $abduction_I = abduction_S \cup induction_S$. $Induction_I$ does not have an analogue in the syllogistic perspective since the process of evaluating a given hypothesis against reality is not reasoning at all. To avoid terminological problems I think that it is better and less confusing to speak about *hypothesis evaluation* or *hypothesis testing*. Furthermore, I suggest to use the term *conjectural reasoning* (or *hypothetical reasoning*) for $abduction_I$, because it puts abduction on an equal footing with induction, and because it conforms to what seems to me the majority practice in AI.

On the other hand, I believe the inferential perspective to be the more fundamental of the two. Here one might draw an analogy with nonmonotonic reasoning, where one has, on the one hand, the inferential perspective of (say) preferential consequence relations, while on the other hand there are the more syntactically oriented approaches such as negation as failure, Closed-World Assumption, Predicate Completion, and what have you. Semantically speaking these approaches employ a particular preference ordering, which cannot be completely axiomatised on the level of consequence relations. On the other hand, the inferential perspective on nonmonotonic reasoning cannot teach us anything about particular preference orderings, but does give fundamental insights into the consequences of using any preference ordering at all. Analogously, conjectural consequence relations can be used to gain deeper insight into the process of hypothesis formation, be they abductive or inductive hypotheses. So we might say that the two perspectives are complementary rather than orthogonal.

From the inferential perspective there may be several intuitions regarding hypotheses: they can be explanatory, accounting for part of the evidence, or confirmatory, extracting regularities implicitly present in the evidence. Failing to discern these intuitions may lead to paradox [11] as the two notions have partly conflicting formal properties. Confirmation is a qualitative generalisation of logical entailment, and so it is perhaps not surprising that the formalisation of confirmatory reasoning presented here is almost identical to the formalisation of nonmonotonic reasoning by means of preferential models, even though some of the intuitions are different. In fact, in the inferential context of confirmatory reasoning one may again draw a syllogistic distinction between inferring default hypotheses ('Tweety flies') and inferring inductive hypotheses ('crows are black'). As with abduction and induction in logic programming the underlying logical mechanisms may be the same, but the methods and pragmatics are quite different.

Finally I mention some related work. Console & Saitta [15] discuss different semantics for abduction and induction, while Dimopoulos & Kakas [16] investigate possible interactions between the two methods. Jennings, Chan & Dowad argue for 'a generalisation of inference from the standard account in terms of truth preservation to one which countenances preservation of other desirable metalinguistic properties' [17], foreshadowing my definition of explanatory consequence as preservation of explanatory power. Another predecessor of this definition is Zadrozny [18].

In this paper I have attempted to clarify abduction and induction by investigating the two perspectives developed by Peirce. The syllogistic perspective is intuitively appealing, but stands in need of a semantic characterisation. The inferential perspective concentrates on the underlying logic of hypothesis formation, for which I have suggested a formal framework in terms of conjectural consequence relations. Future work includes the following: (*i*) semantic characterisation of abduction and induction in a single framework, (*ii*) axiomatisation of explanatory reasoning for weaker explanation mechanisms, and (*iii*) development of a more general form of confirmatory reasoning, not satisfying Right And.

ACKNOWLEDGEMENTS

Part of this work was supported by Esprit Long Term Research Project 20237 (Inductive Logic Programming 2) and PECO Pan-European Scientific Network ILPnet (no. CIPA3510OCT920044). The paper has benefitted from discussions with Marc Denecker, Tony Kakas and Bern Martens.

REFERENCES

[1] A.C. Kakas, R.A. Kowalski & F. Toni, 'Abductive Logic Programming', *J. Logic and Computation*, **2**(6): 719-770 (1992).

[2] J.R. Josephson, 'Conceptual analysis of abduction', 5–30, *Abductive inference: computation, philosophy, technology*, J.R. Josephson & S.G. Josephson (eds), Cambridge University Press, Cambridge, 1994.

[3] R.S. Michalski, R.S., 'Concept learning', 185–194, *Encyclopedia of Artificial Intelligence*, S.C. Shapiro (ed.), John Wiley, Chicester, 1987.

[4] S. Kraus, D. Lehmann & M. Magidor, 'Nonmonotonic reasoning, preferential models and cumulative logics', *Art. Int.* **44**: 167–207 (1990).

[5] C. Harstshorne, P. Weiss & A. Burks, eds., *Collected Papers of Charles Sanders Peirce*, Harvard University, Cambridge, 1931–58.

[6] K.T. Fann, *Peirce's theory of abduction*, Martinus Nijhoff, The Hague, 1970.

[7] R. Carnap, *Logical Foundations of Probability*, Routledge & Kegan Paul, London, 1950.

[8] D. Poole, 'A logical framework for default reasoning', *Art. Int.* **36**: 27-47 (1988).

[9] J.R. Quinlan, 'Learning logical definitions from relations', *Machine Learning* **5**(3): 239–266 (1990).

[10] P.A. Flach, *Conjectures: an inquiry concerning the logic of induction*, PhD thesis, Tilburg University, 1995.

[11] C.G. Hempel, 'Studies in the logic of confirmation', *Mind* **54**(213): 1–26; **54**(214): 97–121 (1945).

[12] N. Helft, 'Induction as nonmonotonic inference', 149–156, *Proc. KR'89*, Morgan Kaufmann, San Mateo, 1989.

[13] P.A. Flach, 'Predicate invention in Inductive Data Engineering', 83–94, *Proc. ECML'93*, P.B. Brazdil (ed.), LNAI 667, Springer-Verlag, Berlin, 1993.

[14] L. De Raedt & M. Bruynooghe, 'A theory of clausal discovery', 1058–1063, *Proc. IJCAI'93*, Morgan Kaufmann, San Mateo, 1993.

[15] L. Console & L. Saitta, 'Generalization and explanation revisited in Machine Learning', manuscript (1994).

[16] Y. Dimopoulos & A. Kakas, 'Abduction and inductive learning', manuscript (1995).

[17] R.E. Jennings, C.W. Chan & M.J. Dowad, 'Generalised inference and inferential modelling', 1046–1051, *Proc. IJCAI'91*, Morgan Kaufmann, San Mateo, 1991.

[18] W. Zadrozny, *On rules of abduction*, IBM Research Report, 1991.