A model of induction

Flach, P.A.

Publication date:
1992

Document Version
Publisher's PDF, also known as Version of record

Link to publication in Tilburg University Research Portal

Citation for published version (APA):
Flach, P. A. (1992). *A model of induction*. (ITK Research Report). Institute for Language Technology and Artifical Intelligence, Tilburg University.

[redacted]

ITK RESEARCH REPORT

INSTITUTE FOR LANGUAGE TECHNOLOGY AND ARTIFICIAL INTELLIGENCE

INSTITUUT VOOR TAAL- EN KENNISTECHNOLOGIE

ITK Research Report
December 21, 1992

A model of induction

Peter A. Flach

No. 41 1992/41

ISSN 0924-7807

Tilburg University, P.O.Box 90153, 5000 LE Tilburg, The Netherlands
Phone: +3113 663113, Fax: +3113 663110.

A model of inductive reasoning[†]

Peter A. Flach

ABSTRACT

In this paper, we formally characterise the process of inductive hypothesis formation. This is achieved by formulating minimal properties for inductive consequence relations. These properties are justified by the fact that they are sufficient to allow identification in the limit. By means of stronger sets of properties, we define both standard and non-standard forms of inductive reasoning, and we give an application of the latter.

Contents

[†] This paper was presented at the Applied Logic Conference "Logic at Work" (Amsterdam, December 17-19, 1992).

1. Introduction

1.1 Motivation and scope

Induction is the process of drawing conclusions about all members of a certain set from knowledge about specific instances of that set. For example, after observing a number of black crows, we might conclude inductively that all crows are black. Such a conclusion can never be drawn with absolute certainty, and an immediate question is: how is our confidence in it affected by observing the next black crow? This problem is known as the *justification problem* of induction, a problem with which philosophers of all times have wrestled without finding a satisfactory solution.

In this paper, we are concerned with a different but related problem: the formalisation of the process of *inductive hypothesis formation*. Which hypotheses are possible, given the available information? For instance, in the crows example the hypothesis 'all crows are black' is possible, but the hypothesis 'all crows are white' is not: it is refuted as soon as we observe one black crow. Moreover, once refuted, it will never become a possible hypothesis again, no matter how many crows are observed. The question is thus: what is the relation between sets of observations and possible hypotheses?

In order to address this question, we need a representation for observations and hypotheses. In our framework, they are represented by logical statements. Because of the distinction between instances and sets, a first-order logic seems most appropriate. However, this is not mandatory, as a simple example will show. Consider the problem of acquiring a concept from observed instances. Simple concepts and their instances might be described in a socalled *attribute-value language*, expressing knowledge about properties like colour, shape and size. Attribute-value languages have the expressive power of propositional logic. It follows that the distinction between instances and sets is not essential for performing induction[1].

Since observations and hypotheses are logical statements, inductive hypothesis formation can be declaratively modeled as a consequence relation. We will study the properties of such inductive consequence relations, thereby applying techniques developed in other fields of non-standard logic, like non-monotonic reasoning (Gabbay, 1985; Shoham, 1987; Kraus, Lehmann & Magidor, 1990), abduction (Zadrozny, 1991), and belief revision (Gärdenfors, 1988; 1990). In the spirit of these works, we develop several systems of properties inductive consequence relations might have. The results are twofold: our weakest system **I** gives *necessary conditions* for inductive consequence relations, and the other systems are grouped together in two main families, modeling *different kinds* of induction.

[1]For simplicity, the framework in this paper will be developed on the basis of propositional logic; we have no reason to believe, however, that its applicability cannot be extended to first-order logic.

As a motivating example for this latter point, let the observations be drawn from a database of facts about different persons, including their first and last names, and their parents' first and last names. A typical inductive hypothesis would be 'every person's last name equals her fathers' last name'. Such a hypothesis, if adopted, would yield a procedure for finding a person's last name, given her father's last name. Now consider the statement 'every person has exactly one mother'. It is also an inductive hypothesis, but of a different kind. Specifically, it does not give a procedure for finding a person's mother (such a procedure does obviously not exist), but merely states her existence and uniqueness. While the first hypothesis can be seen as a *definition* of the last name of children, the second is instead a *constraint* on possible models of the database. In the sections to follow, we will give practical examples of both forms of induction.

1.2 Terminology and notation

Suppose P is a computer program that performs inductive reasoning. That is, P takes a set of formulas α in some language L as input, and outputs inductive conclusions β. The main idea is to view P as constituting a consequence relation, i.e. a relation on $2^L \times L$, and to study the properties of this consequence relation. We will write $\alpha \mathrel{|\!<} \beta$ whenever β is an inductive consequence of the premises α. A set of premises is often represented by a formula expressing their conjunction. The properties of $\mathrel{|\!<}$ will be expressed by Gentzen-style inference rules in a meta-logic, following (Gabbay, 1985; Kraus, Lehmann & Magidor, 1990).

We will assume that L is a propositional language, closed under the logical connectives. Furthermore, we assume a set of models M for L, and the classical satisfaction relation $\models$ on $M \times L$. If $m \models \alpha$ for all $m \in M$, we write $\models \alpha$. We can implicitly introduce background knowledge by restricting M to a proper subset of all possible models. We will assume that $\models$ is compact, i.e. an infinite set of formulas is unsatisfiable if and only if every finite subset is.

In many practical cases premises and hypotheses are drawn from restricted sublanguages of L. Given a language L, an *inductive frame* is a triple $\langle \Gamma, \mathrel{|\!<}, \Sigma \rangle$, where $\Gamma \subseteq L$ is the set of possible observations, $\Sigma \subseteq L$ is the set of possible hypotheses, and $\mathrel{|\!<}$ is an inductive consequence relation on $2^L \times L$. We will assume that Γ is at least closed under conjunction.

The fact that in an inductive frame the consequence relation is defined on $2^L \times L$, rather than $2^\Gamma \times \Sigma$, reflects an important choice for a certain interpretation of $\mathrel{|\!<}$. Specifically, we chose to interpret $\alpha \mathrel{|\!<} \beta$ not just as 'β is an inductive consequence of α', but more generally as 'β is a possible hypothesis, given α'. In this way, our framework allows the study of not only inductive reasoning, but hypothetical reasoning in general.

As an example, take the property of Contraposition (if $\alpha \mathrel{|\!<} \beta$, then $\neg\beta \mathrel{|\!<} \neg\alpha$), which we will encounter later. This property can be understood as follows. Suppose I know that c is a crow (background knowledge), and the premiss that c is black (α) allows the inductive hypothesis 'all crows are black' (β), then the premiss that some crows are not black ($\neg\beta$) allows likewise the hypothesis that c is not black

($\neg\alpha$). It is quite possible that, in most inductive frames, this hypothesis is excluded from Σ[2]. This prohibits the interpretation of Contraposition as stating 'if $\alpha \mathrel{|\!\!<} \beta$ is an inductive argument, then so is $\neg\beta \mathrel{|\!\!<} \neg\alpha$'. Rather, it describes a property of hypothesis formation in general: '$\neg\alpha$ is a possible hypothesis on the basis of $\neg\beta$, just like β is a possible hypothesis on the basis of α'.

The plan of the paper is as follows. In section 2, we define a minimal set of properties for inductive consequence relations, and we show that these properties are sufficient in the sense that they allow for a very general induction method. In sections 3 and 4, we develop two, more or less complementary, kinds of inductive reasoning, and we give their main properties. We end the paper with some concluding remarks.

2. Identification in the limit and I-relations

2.1 Identification in the limit

Inductive arguments are *defeasible*: an inductive conclusion might be invalidated by future observations. Thus, the validity of an inductive argument can only be guaranteed when complete information is available. A possible way to model complete information is by a sequence of formulas (possibly infinite), such that every incorrect hypothesis is eventually ruled out by a formula in the sequence. If an inductive reasoner reads in a finite initial segment of this sequence and outputs a correct hypothesis, it is said to have *finitely identified* the hypothesis. Since this is a fairly strong criterion, it is often weakened as follows: the inductive reasoner is allowed to output as many hypotheses as wanted, but after finitely many guesses the hypothesis must be correct, and not abandoned afterwards. This is called *identification in the limit*. The difference with finite identification is, that the inductive reasoner does not know when the correct hypothesis has been attained. Details can be found in (Gold, 1967).

We will redefine identification in the limit in terms of inductive consequence relations. Given a set of hypotheses Σ, the task is to identify an unknown $\beta \in \Sigma$ from a sequence of observations $\alpha_1, \alpha_2, \ldots$, such that $\{\alpha_1, \alpha_2, \ldots\} \mathrel{|\!\!<} \beta$. The observations must be sufficient in the sense that they eventually rule out every non-intended hypothesis.

DEFINITION 2.1. *Identification in the limit.*

Let $<\Gamma, \mathrel{|\!\!<}, \Sigma>$ be an inductive frame. Given a *target hypothesis* $\beta \in \Sigma$, a *presentation for β* is a (possibly infinite) sequence of observations $\alpha_1, \alpha_2, \ldots$ such that $\{\alpha_1, \alpha_2, \ldots\} \mathrel{|\!\!<} \beta$. Given a presentation, an *identification algorithm* is an algorithm which reads in an observation α_j from the presentation and outputs a hypothesis β_j, for j=1, 2, ... The output sequence $\beta_1, \beta_2, \ldots$ is said to *converge* to β_n if for all $k \geq n$, $\beta_k = \beta_n$.

[2]In fact, it looks more like an abductive hypothesis. We have argued elsewhere (Flach, 1992) that there are close relations between abduction and induction.

A presentation $\alpha_1,\alpha_2,\ldots$ for β is *sufficient* if for any hypothesis $\gamma\in\Sigma$ other than β it contains a *witness* α_i such that $\{\alpha_1,\alpha_2,\ldots,\alpha_i\} \not\mathrel{\kappa} \gamma$. An identification algorithm is said to *identify β in the limit* if, given any sufficient presentation for β, the output sequence converges to β. An identification algorithm identifies Σ in the limit, if it is able to identify any $\beta\in\Sigma$ in the limit.

Since we place induction in a logical context, it makes sense not to distinguish between logically equivalent hypotheses. That is, β is *logically* identified in the limit if the output sequence converges to β' such that $\models\beta'\leftrightarrow\beta$. A presentation for β is *logically* sufficient if it contains a witness for any γ such that $\not\models\gamma\leftrightarrow\beta$. In the sequel, we will only consider logical identification, and omit the adjective 'logical'.

2.2 I-relations

After having defined identification in the limit in terms of inductive consequence relations, we now turn to the question: what does it take for inductive consequence relations to behave sensibly? We will first consider some useful properties, and then combine these properties into formal system **I**. We consider this system to be the weakest possible system defining inductive consequence relations.

The first two properties follow from the definition of identification in the limit. Suppose that α_i is a witness for γ, i.e. $\{\alpha_1,\alpha_2,\ldots,\alpha_i\} \not\mathrel{\kappa} \gamma$, then any extended set of observations should still refute γ, i.e. $\{\alpha_1,\alpha_2,\ldots,\alpha_i\}\cup A \not\mathrel{\kappa} \gamma$ for any $A\subseteq\Gamma$. Conversely, if $B \mathrel{\kappa} \beta$ then also $B' \mathrel{\kappa} \beta$ for any $B'\subseteq B$. Assuming that sets of observations can always be represented by their conjunction, this property can be stated as follows:

(1) $$\frac{\models\alpha\rightarrow\beta\ ,\ \alpha \mathrel{\kappa} \gamma}{\beta \mathrel{\kappa} \gamma}$$

Furthermore, observations can not distinguish between logically equivalent hypotheses:

(2) $$\frac{\models\beta\leftrightarrow\gamma\ ,\ \alpha \mathrel{\kappa} \beta}{\alpha \mathrel{\kappa} \gamma}$$

The other two properties are not derived directly from identification in the limit. Instead, they describe the relation between observations and the hypotheses they confirm or refute. Here, the basic assumption is that induction aims to increase knowledge about some unknown intended model m_0. The observations are obtained from a reliable source, and are therefore true in m_0. On the other hand, hypotheses represent assumptions about the intended model. Together, observations and hypotheses can be used to make predictions about m_0. More specifically, suppose we have adopted hypothesis β on the basis of observations α, and let δ be a logical consequence of $\alpha\wedge\beta$, then we expect δ to be true in m_0. If the next observation conforms to our prediction, then we stick to β; if it contradicts our prediction, β should be refuted. These two principles can be expressed as follows.

(3) $$\frac{\models\alpha\wedge\beta\rightarrow\delta\ ,\ \alpha \mathrel{\kappa} \beta}{\alpha\wedge\delta \mathrel{\kappa} \beta}$$

(4) $$\frac{\models\alpha\wedge\beta\rightarrow\delta\ ,\ \alpha \mathrel{\kappa} \beta}{\alpha\wedge\neg\delta \not\mathrel{\kappa} \beta}$$

Note that the combination of these rules requires that $\alpha\wedge\beta$ is consistent: otherwise, we would have both $\models\alpha\wedge\beta\rightarrow\delta$ and $\models\alpha\wedge\beta\rightarrow\neg\delta$, and thus both $\alpha\wedge\delta \mathrel{|\!\!<} \beta$ (3) and $\alpha\wedge\delta \mathrel{|\!\!\not<} \beta$ (4). For technical reasons, the inconsistency of $\alpha\wedge\beta$ is not prohibited a priori. In the presence of the other rules, the application of rule (4) can be blocked in this case by adding the consistency of β as a premiss (theorem 2.4).

Rules (1) and (3) look pretty similar, and can probably be combined into a single rule. Rule (2) is clearly independent from the other rules. Rule (4) is not derivable from the other rules, but it may be if we add a weaker version. These considerations lead to the following system of rules.

DEFINITION 2.2. *I-relations.*

The system **I** consists of the following four rules:

- **Conditional Reflexivity:** $$\frac{\not\models\neg\alpha}{\alpha \mathrel{|\!\!<} \alpha}$$
- **Consistency:** $$\frac{\not\models\neg\alpha}{\neg\alpha \mathrel{|\!\!\not<} \alpha}$$
- **Right Logical Equivalence:** $$\frac{\models\beta\leftrightarrow\gamma\ ,\ \alpha \mathrel{|\!\!<} \beta}{\alpha \mathrel{|\!\!<} \gamma}$$
- **Convergence:** $$\frac{\models\alpha\wedge\gamma\rightarrow\beta\ ,\ \alpha \mathrel{|\!\!<} \gamma}{\beta \mathrel{|\!\!<} \gamma}$$

If $\mathrel{|\!\!<}$ is a consequence relation satisfying the rules of *I*, it is called an *I-relation.*

The following lemma gives two useful derived rules in this system.

LEMMA 2.3. *The following rules are derived rules in system* ***I****:*

- **S** $$\frac{\not\models\neg\beta\ ,\models\beta\rightarrow\alpha}{\alpha \mathrel{|\!\!<} \beta}$$
- **W** $$\frac{\not\models\neg\beta\ ,\ \alpha \mathrel{|\!\!<} \beta}{\not\models\beta\rightarrow\neg\alpha}$$

Proof. (S) Suppose $\not\models\neg\beta$ and $\models\beta\rightarrow\alpha$; by Conditional Reflexivity it follows that $\beta \mathrel{|\!\!<} \beta$, and we conclude by Convergence.

(W) Suppose $\not\models\neg\beta$ and $\alpha \mathrel{|\!\!<} \beta$; by Consistency it follows that $\neg\beta \mathrel{|\!\!\not<} \beta$. By Convergence, it follows that $\not\models\alpha\wedge\beta\rightarrow\neg\beta$, i.e. $\not\models\beta\rightarrow\neg\alpha$. ■

Rule S expresses that β is an inductive consequence of α if α is deductively entailed by β; rule W states that β shouldn't deductively entail $\neg\alpha$ if it is an inductive consequence of α (β consistent in both cases). Rule S hints at the view of induction as reversed deduction, while rule W suggests a connection between induction and finding consistent extensions of a theory. By strengthening one of these rules, we obtain systems for one of two different kinds of induction: strong induction (section 3) and weak induction (section 4).

The following theorem shows that system **I** does what it was intended to do.

THEOREM 2.4. *Rules (1)—(4) are derived rules in system* ***I****.*

Proof. (1) Suppose $\models\alpha\rightarrow\beta$, i.e. $\models\alpha\wedge\gamma\rightarrow\beta$ and $\alpha \mathrel{|\!\!<} \gamma$; we have $\alpha \mathrel{|\!\!<} \beta$ by Convergence.

(2) Identical to Right Logical Equivalence.

(3) Suppose $\models\alpha\wedge\beta\rightarrow\delta$, i.e. $\models\alpha\wedge\beta\rightarrow\alpha\wedge\delta$ and $\alpha \mathrel{|\!\!<} \beta$; by Convergence it follows that $\alpha\wedge\delta \mathrel{|\!\!<} \beta$. Note that in the presence of rule (1), rule (3) is equivalent to Convergence, since the latter can also be derived from the former: suppose $\models\alpha\wedge\gamma\rightarrow\beta$ and $\alpha \mathrel{|\!\!<} \gamma$, then by (3) $\alpha\wedge\beta \mathrel{|\!\!<} \gamma$, and since $\models\alpha\wedge\beta\rightarrow\beta$, by (1) $\beta \mathrel{|\!\!<} \gamma$. Since we already showed that (1) follows from Convergence, we conclude that Convergence exactly replaces (1) and (3).

(4) As said earlier, we prove this rule under the assumption that β is consistent. Suppose $\alpha\wedge\neg\delta \mathrel{|\!\!<} \beta$, then by rule W $\not\models\beta\rightarrow\neg(\alpha\wedge\neg\delta)$, i.e. $\not\models\alpha\wedge\beta\rightarrow\delta$. From $\alpha \mathrel{|\!\!<} \beta$ and the consistency of β, it follows that $\alpha\wedge\beta$ is consistent by rule W, as required. ■

It should be noted that Conditional Reflexivity is nowhere used in the proof of theorem 2.4. This indicates that it can be removed to obtain a truly minimal rule system for induction. However, system **I** possesses a nice symmetry, as shown by the next result.

THEOREM 2.5. *Define* $\alpha \mathrel{|\!\!<\!\!<} \beta$ *iff* $\neg\alpha \not\mathrel{|\!\!<} \beta$*, then* $\mathrel{|\!\!<}$ *is an I-relation iff* $\mathrel{|\!\!<\!\!<}$ *is an I-relation.*

Proof. Using the rewrite rule $\alpha \mathrel{|\!\!<} \beta \Rightarrow \neg\alpha \not\mathrel{|\!\!<\!\!<} \beta$, Conditional Reflexivity rewrites to Consistency and *vice versa*, while Right Logical Equivalence and Convergence rewrite to themselves. Since this rewrite rule is its own inverse, this proves the theorem in both directions. ■

This duality will reappear later, as it provides the link between weak and strong induction (section 4.2).

System **I** has been built on the basis of rules (1)—(4), which in turn were derived from the notion of identification in the limit. The following section rounds off this analysis by demonstrating how one could use inductive consequence relations for performing the perhaps most elementary form of identification: identification by enumeration.

2.3 Identification by enumeration

If we assume that the set of inductive hypotheses is countable, we can formulate a very simple and general identification algorithm (Algorithm 2.6). We enumerate all the possible hypotheses, and search this enumeration for a hypothesis that is an inductive consequence of the premises seen so far. We stick to this hypothesis until we encounter a new premise which, together with the previous premises, contradicts it: then we continue searching the enumeration.

ALGORITHM 2.6. *Identification by enumeration.*
Input: a presentation $\alpha_1,\alpha_2,\ldots$ for a target hypothesis $\beta\in\Sigma$, and an enumeration $\beta_1,\beta_2,\ldots$ of all the formulas in Σ.
Output: a sequence of formulas in Σ.

```
begin
        i:=1; k:=1;
        repeat
                while {α_j | j≤i} ⊭ β_k do k:=k+1;
                output β_k;
                i:=i+1;
        forever;
end.
```

Algorithm 2.6 is very powerful, but it has one serious drawback: the enumeration of hypotheses is completely unordered. Therefore, there is much duplication of work in checking hypotheses. There exist more practical versions of this algorithm, that can be applied if the set of hypotheses can be ordered. However, it is clear that if any search-based identification algorithm can achieve identification in the limit, identification by enumeration can, provided the inductive consequence relation is 'well-behaved'. The following theorem states that I-relations are well-behaved in this sense.

THEOREM 2.7. *Algorithm 2.6 performs identification in the limit if* $\succ$ *is an I-relation.*

Proof. Let α denote the entire presentation, and let β be the target hypothesis, i.e. $\alpha \succ \beta$. Furthermore, let β_n be the first formula in the enumeration, such that $\models\beta_n\leftrightarrow\beta$. We will show that the output sequence converges to β_n if α is sufficient for β.

Suppose β_k, $k<n$ precedes β_n in the presentation. By assumption, $\not\models\beta_k\leftrightarrow\beta$; if the presentation is sufficient, there will be a witness α_i such that $\{\alpha_1,\alpha_2,\ldots,\alpha_i\} \not\succ \beta_k$, so β_k will be discarded.

Since $\models\beta_n\leftrightarrow\beta$ and $\alpha \succ \beta$, it follows by Right Logical Equivalence that $\alpha \succ \beta_n$. By Convergence, $\alpha' \succ \beta_n$ for every initial segment α'. Therefore β_n is never discarded. ■

Note that this proof only mentions the rules Right Logical Equivalence and Convergence. As said before, Consistency is needed to ensure that the presentation and the hypothesis can be combined in a meaningful way, and Conditional Reflexivity is not strictly needed.

Induction would be infeasible if it could only be achieved by enumerating hypotheses. Likewise, inductive consequence relations would be useless if nothing stronger than I-relations would exist. In the following two sections, we present two families of inductive consequence relations, the first based on a view of induction as reversed deduction, and the second based on a view of induction as trying to extend the observations consistently. We will thereby adopt a terminology that is more familiar in the field of Machine Learning: observations are called *examples*, and inductive hypotheses are also called *explanations* of the examples.

3. Strong induction

A *strong* inductive consequence relation is an I-relation that satisfies the following rule:

- S′ $\dfrac{\models\beta\rightarrow\alpha}{\alpha \mathrel{|\!\!<} \beta}$

This is a strengthening of the derived rule S in system I (lemma 2.3). The idea of strong induction is, that it equals reversed deduction in some underlying logic or *base logic*. Rule S′ states, that this base logic should allow all valid classical deductions. The base logic might also allow deductions that are classically invalid, but (for instance) plausible. Rule S′ is a strengthening of S, because it doesn't require β to be consistent. In general, inconsistent inductive hypotheses are not very interesting; they arise as a borderline case, similar to tautologies that are deductive consequences of any set of premises. In the present context, this borderline case is instrumental in distinguishing strong induction from weak induction, as we will see in section 4.

Rule S′ can be derived if we strengthen Conditional Reflexivity to

- **Reflexivity:** $\alpha \mathrel{|\!\!<} \alpha$

Thus, an I-relation is a strong inductive consequence relation iff it satisfies Reflexivity.

3.1 The system SC

The weakest system of rules for strong induction is called **SC**, which stands for Strong induction with a *Cumulative* base logic. For inductive consequence relations, cumulativity means that if $\beta \mathrel{|\!\!<} \gamma$, the hypotheses γ and $\beta\wedge\gamma$ inductively explain exactly the same facts. This principle can be expressed by two rules: Right Cut and Right Extension.

DEFINITION 3.1. *The system SC.*

The system SC consists of the following six rules:

- **Reflexivity:** $\alpha \mathrel{|\!\!<} \alpha$
- **Consistency:** $\dfrac{\not\models\neg\alpha}{\neg\alpha \mathrel{|\!\!\not<} \alpha}$
- **Right Logical Equivalence:** $\dfrac{\models\beta\leftrightarrow\gamma\ ,\ \alpha \mathrel{|\!\!<} \beta}{\alpha \mathrel{|\!\!<} \gamma}$
- **Convergence:** $\dfrac{\models\alpha\wedge\gamma\rightarrow\beta\ ,\ \alpha \mathrel{|\!\!<} \gamma}{\beta \mathrel{|\!\!<} \gamma}$
- **Right Cut:** $\dfrac{\alpha \mathrel{|\!\!<} \beta\wedge\gamma\ ,\ \beta \mathrel{|\!\!<} \gamma}{\alpha \mathrel{|\!\!<} \gamma}$
- **Right Extension:** $\dfrac{\alpha \mathrel{|\!\!<} \gamma\ ,\ \beta \mathrel{|\!\!<} \gamma}{\alpha \mathrel{|\!\!<} \beta\wedge\gamma}$

Right Cut expresses that a part of an inductive hypothesis, which inductively explains another part, may be cut away from the hypothesis. In the presence of rule S′, it is a strengthening of Convergence. Right Extension states that an inductive hypothesis may be extended by some of the things it explains.

These latter two rules may look suspicious, because β takes the role of both example and hypothesis. For instance, Right Extension might not be applicable in a particular inductive frame, because $\beta \notin \Sigma$. The reader will recall the discussion in section 1.2, where it was argued that even if this is so, such rules may describe useful properties of the process of (inductive) hypothesis formation. Here we encounter a case in point, because the two new rules interact to produce a rule that is satisfied in any inductive frame of which the consequence relation satisfies the rules of **SC**.

LEMMA 3.2. *In* ***SC****, the following rule can be derived:*

- **Compositionality:** $$\frac{\alpha \Bbbk \gamma\ ,\ \beta \Bbbk \gamma}{\alpha \wedge \beta \Bbbk \gamma}$$

Proof. Suppose $\alpha \Bbbk \gamma$ and $\beta \Bbbk \gamma$; by Right Extension we have $\alpha \Bbbk \beta \wedge \gamma$. Also, because $\alpha \wedge \beta \wedge \gamma \models \alpha \wedge \beta$, we have $\alpha \wedge \beta \Bbbk \alpha \wedge \beta \wedge \gamma$ by rule S′. Using Right Cut gives $\alpha \wedge \beta \Bbbk \beta \wedge \gamma$, and since by assumption $\beta \Bbbk \gamma$, we can cut away β from the righthand side to get $\alpha \wedge \beta \Bbbk \gamma$. ■

Compositionality states that if an inductive hypothesis explains two examples separately, it also explains them jointly. It can be employed to speed up enumerative identification algorithms. Recall that in Algorithm 2.6 a new hypothesis must be checked against the complete set of previously seen examples. If we already know that the new hypothesis inductively explains some subset of those examples, then by Compositionality the remaining examples can be tested in isolation.

Furthermore, if the search strategy guarantees that the new hypothesis explains all the examples explained by the previous hypothesis, then we only need to test it against the last example which refuted the previous hypothesis. This requires an ordering of the hypothesis space, which in turn requires monotonicity of the base logic. This results in the following stronger system.

3.2 The system SM

There are several ways to define monotonicity of the base logic, for instance by adopting transitivity or contraposition.

DEFINITION 3.3. *The system* ***SM****.*

The system **SM** consists of the rules of **SC** plus the following rule:

- **Contraposition:** $$\frac{\alpha \Bbbk \beta}{\neg\beta \Bbbk \neg\alpha}$$

As the following lemma shows, this results in a considerably more powerful system.

LEMMA 3.4. *In SM, the following rules can be derived:*

- **Explanation Strengthening:** $$\frac{\models \gamma \rightarrow \beta \ , \ \alpha \mathrel{|\!\!<} \beta}{\alpha \mathrel{|\!\!<} \gamma}$$

- **Explanation Updating:** $$\frac{\models \gamma' \rightarrow \gamma \ , \ \alpha \mathrel{|\!\!<} \gamma \ , \ \beta \mathrel{|\!\!<} \gamma'}{\alpha \wedge \beta \mathrel{|\!\!<} \gamma'}$$

Proof. (Explanation Strengthening) Suppose $\models \gamma \rightarrow \beta$ and $\alpha \mathrel{|\!\!<} \beta$; by Contraposition, it follows that $\neg\beta \mathrel{|\!\!<} \neg\alpha$. Convergence gives $\neg\gamma \mathrel{|\!\!<} \neg\alpha$, which finally results in $\alpha \mathrel{|\!\!<} \gamma$ by Contraposition.

(Explanation Updating) Suppose $\models \gamma' \rightarrow \gamma$ and $\alpha \mathrel{|\!\!<} \gamma$; by Explanation Strengthening we have $\alpha \mathrel{|\!\!<} \gamma'$. Assuming $\beta \mathrel{|\!\!<} \gamma'$, this gives $\alpha \wedge \beta \mathrel{|\!\!<} \gamma'$ by Compositionality. ■

Explanation Strengthening expresses that any γ logically implying some inductive explanation β of a set of examples α is also an explanation of α. Consequently, the set of inductive explanations of a given set of examples is completely determined by its weakest elements according to logical implication. Since logical implication is reflexive and transitive, it is a quasi-ordering on Σ, which can be turned into a partial ordering by considering equivalence classes of logically equivalent formulas (in other words, the Lindenbaum algebra of Σ).

Explanation Updating is a combination of Explanation Strengthening and Compositionality, which shows how to employ this ordering in identification algorithms. It states that if γ is a hypothesis explaining the examples seen so far α but not the next example β, it can be replaced by some γ' which (*i*) logically implies γ and (*ii*) explains β. This clearly shows that we don't need to test the new hypothesis γ' against the previous examples α.

The properties expressed by these rules have been used in many AI-approaches to inductive reasoning (Mitchell, 1982; Shapiro, 1983). The results in this section have been presented to show how they can be derived systematically within our framework. For instance, we have shown that an important property like Explanation Strengthening requires monotonicity of the base logic.

4. Weak induction

The ideas described in this section have been the main motivating force for the research reported in this paper. While induction and deduction are closely related, they can be related in more than one way. Weak induction provides an alternative for strong induction, which only considers inductive hypotheses from which the examples are provable. Weak induction aims at supplementing the examples with knowledge which is only implicitly contained in those examples.

In section 1.1, we provided an example of weak induction: inferring 'every person has exactly one mother' from a collection of facts. This is not a strong inductive argument, since the induced rule does not entail the facts (regardless of the base logic). Rather, the rule does not contradict the facts: it should not entail their negation. That is, weak inductive consequence relations satisfy the following rule:

- **W′** $\dfrac{\alpha \mathrel{|\!\!<} \beta}{\not\models \beta \rightarrow \neg\alpha}$

Note that this disallows the possibility that β is inconsistent, showing that some strong inductive consequence relations are not weak inductive consequence relations.

Rule W′ is a strengthening of rule W, that can be obtained by strengthening Consistency to

- **Weak Reflexivity:** $\neg\alpha \not\mathrel{|\!\!<} \alpha$

Weak Reflexivity expresses that an inductive hypothesis never explains its negation.

4.1 The system WC

The weakest system for weak inductive reasoning is called **WC**. It models weak induction with a cumulative base logic. The principle of cumulativity for weak inductive consequence relations is stated as follows: if $\neg\beta \not\mathrel{|\!\!<} \gamma$, then β can be added to the inductive hypothesis γ without changing the set of examples it explains. This principle requires weak counterparts of the corresponding rules in **SC**.

DEFINITION 4.1. *The system **WC**.*

The system **WC** consists of the following six rules:

- **Conditional Reflexivity:** $\dfrac{\not\models \neg\alpha}{\alpha \mathrel{|\!\!<} \alpha}$
- **Weak Reflexivity:** $\neg\alpha \not\mathrel{|\!\!<} \alpha$
- **Right Logical Equivalence:** $\dfrac{\models \beta \leftrightarrow \gamma \;,\; \alpha \mathrel{|\!\!<} \beta}{\alpha \mathrel{|\!\!<} \gamma}$
- **Convergence:** $\dfrac{\models \alpha \wedge \gamma \rightarrow \beta \;,\; \alpha \mathrel{|\!\!<} \gamma}{\beta \mathrel{|\!\!<} \gamma}$
- **Weak Right Cut:** $\dfrac{\alpha \mathrel{|\!\!<} \beta \wedge \gamma \;,\; \neg\beta \not\mathrel{|\!\!<} \gamma}{\alpha \mathrel{|\!\!<} \gamma}$
- **Weak Right Extension:** $\dfrac{\alpha \mathrel{|\!\!<} \gamma \;,\; \neg\beta \not\mathrel{|\!\!<} \gamma}{\alpha \mathrel{|\!\!<} \beta \wedge \gamma}$

In this system, we could derive a weak counterpart to Compositionality, expressing that an example, of which the negation is not explained, can be added to the premises. However, this does not express a very useful property. In general, Compositionality itself does not apply to weak inductive reasoning. Consequently, we must always store all previously seen examples, and check them each time we switch to a new inductive hypothesis (an illustration of this will be provided in section 4.3).

4.2 The system WM

In a monotonic base logic, β does not entail $\neg\alpha$ if and only if α does not entail $\neg\beta$. This property means that a weak inductive consequence relation based on a monotonic base logic is *symmetric*. Again, we stress that although Symmetry is obviously not a property of any form of inductive reasoning, it may be a useful property of the inductive consequence relation involved in weak inductive reasoning.

DEFINITION 4.2. *The system* ***WM***.

The system **WM** consists of the rules of **WC** plus the following rule:

- **Symmetry:** $$\frac{\alpha \mathrel{|\!\!<} \beta}{\beta \mathrel{|\!\!<} \alpha}$$

Similar to **SM**, **WM** induces an ordering of the hypothesis space that can be exploited in enumerative identification algorithms. Search will however proceed in the opposite direction of logically weaker formulas.

LEMMA 4.3. *In* ***WM****, the following rule can be derived:*

- **Explanation Weakening:** $$\frac{\models \beta \rightarrow \gamma\ ,\ \alpha \mathrel{|\!\!<} \beta}{\alpha \mathrel{|\!\!<} \gamma}$$

Proof. Suppose $\models\beta\rightarrow\gamma$ and $\alpha \mathrel{|\!\!<} \beta$; by Symmetry, it follows that $\beta \mathrel{|\!\!<} \alpha$. Convergence gives $\gamma \mathrel{|\!\!<} \alpha$, which finally results in $\alpha \mathrel{|\!\!<} \gamma$ by Symmetry. ■

In fact, **SM** and **WM** are interdefinable in the following sense.

LEMMA 4.4. *Define* $\alpha \mathrel{|\!\!<}\!< \beta$ *iff* $\neg\alpha \not\mathrel{|\!\!<} \beta$, *then* $\mathrel{|\!\!<}$ *satisfies the rules of* ***SM*** *iff* $\mathrel{|\!\!<}\!<$ *satisfies the rules of* ***WM***.

Proof. Using the rewrite rule $\alpha \mathrel{|\!\!<} \beta \Rightarrow \neg\alpha \not\mathrel{|\!\!<}\!< \beta$, each rule of **SM** rewrites (after re-arranging) to a rule of **WM**: Reflexivity rewrites to Weak Reflexivity, Consistency rewrites to Conditional Reflexivity, Convergence and Right Logical Equivalence rewrite to themselves, Right Cut to Weak Right Extension, Right Extension to Weak Right Cut, and Contraposition rewrites to Symmetry. ■

We encountered this transformation before, when we noted that it leaves system **I** invariant (theorem 2.5).

4.3 An application of weak induction

In this section, we will illustrate the usefulness of weak induction by applying it to the problem of inducing integrity constraints in a deductive database. The induction algorithm is fully described in (Flach, 1990), and has also been implemented. Tuples of a database relation (i.e., ground facts) play the role of examples, and hypotheses are integrity constraints on this relation. In the current implementation, hypotheses are restricted to functional and multivalued dependencies between attributes.

Let `child` be a relation with five attributes: child's first name, father's first and last name, and mother's first and last name. Of this relation, the following tuples are given.

```
child(john,frank,johnson,mary,peterson).
child(peter,frank,johnson,mary,peterson).
child(john,robert,miller,gwen,mcintyre).
child(ann,john,miller,dolly,parton).
child(millie,frank,miller,dolly,mcintyre).
```

Table 1. A database relation.

Suppose that we are interested in the attributes that functionally determine the mother's last name (a socalled *functional dependency*). Two such dependencies that are satisfied in table 1 are:

`child(N,_,FL,_,ML1)` ∧ `child(N,_,FL,_,ML2)` → `ML1=ML2`

`child(_,FF,FL,_,ML1)` ∧ `child(_,FF,FL,_,ML2)` → `ML1=ML2`

(we follow the Prolog conventions: all variables are universally quantified, and the underscores denote unique variables). The first formula states that child's first name and father's last name determine mother's last name, and the second formula says that father's first and last names determine mother's last name. Note that these formulas are not logical consequences of the tuples, nor are the tuples logical consequences of the formulas.

How would we induce these dependencies? According to Explanation Weakening, we can start with the strongest hypothesis: all mothers have the same last name (it is determined by the empty set of attributes). This is expressed by the following formula:

`child(_,_,_,_,ML1)` ∧ `child(_,_,_,_,ML2)` → `ML1=ML2`

Since this formula is inconsistent with the tuples in table 1, we will make minimal changes in order to get weaker constraints, which is done by unifying variables on the lefthand side.

For instance, the first and third tuple lead to the following false formula:

`child(john,frank,johnson,mary,peterson)` ∧
`child(john,robert,miller,gwen,mcintyre)` → `peterson=mcintyre`

The formula is false because = is interpreted as syntactical identity. It shows how we can make minimal changes to the original formula: by unifying variables in those positions for which the tuples have different values. This leads to the following three hypotheses:

`child(_,FF,_,_,ML1)` ∧ `child(_,FF,_,_,ML2)` → `ML1=ML2`

`child(_,_,FL,_,ML1)` ∧ `child(_,_,FL,_,ML2)` → `ML1=ML2`

`child(_,_,_,MF,ML1)` ∧ `child(_,_,_,MF,ML2)` → `ML1=ML2`

Each of these hypotheses is again tested for consistency with the data.

If we search in a breadth-first fashion, we will eventually encounter all sets of attributes that determine the mother's last name. Note that, any time we switch to a new hypothesis, we have to check it against the *complete* set of tuples (Compositionality does not hold).

In this setting, there are rather strong restrictions on both Γ (the tuples) and Σ (the functional dependencies). They are needed to ensure convergence of the induction process, and also block properties like Symmetry. On the other hand, Γ should be rich enough to allow sufficient presentations for any hypothesis in Σ (they should form what Shapiro (1983) calls an *admissible pair*). For instance, let Σ be the set of *multivalued dependencies* that hold for a given database relation. An example of such a dependency is

`child(N1,FF1,FL1,MF,ML)` $\wedge$ `child(N2,FF2,FL2,MF,ML)` $\rightarrow$
`child(N1,FF2,FL2,MF,ML)`

which states that children have all the fathers of any child of a certain mother. Such dependencies can be learned in exactly the same way as functional dependencies. The point is, that Γ should now contain positive *and negative* ground facts, since a given multivalued dependency can only be refuted by two tuples in the relation and one tuple known to be not in the relation.

5. Conclusion and future work

The contributions presented in this paper are twofold. First of all, we have given minimal conditions for inductive consequence relations, which are powerful enough to allow identification in the limit. On the other hand, these conditions are liberal enough as to leave room for 'non-standard' forms of inductive reasoning. Our second contribution lies in identifying weak induction as such a non-standard form of induction. We have illustrated the usefulness of weak induction by an example.

As it stands, the framework is far from complete. In particular, a model-theoretic account of induction should accompany our proof-theoretic characterisation. Furthermore, we could study induction with respect to other base logics, such as modal, temporal and intuitionistic logics. Finally, we could investigate system **I** in order to see whether it leaves room for yet another type of induction.

Acknowledgements

I would like to thank John-Jules Meyer, Yao-Hua Tan, Cees Witteveen, and two anonymous referees of the conference 'Logic at Work' for their insightful remarks.

References

P.A. FLACH (1990), 'Inductive characterisation of database relations'. In *Proc. International Symposium on Methodologies for Intelligent Systems*, Z.W. Ras, M. Zemankowa & M.L. Emrich (eds.), pp. 371-378, North-Holland, Amsterdam. Full version appeared as ITK Research Report no. 23.

P.A. FLACH (1992), 'An analysis of various forms of 'jumping to conclusions' '. In *Analogical and Inductive Inference AII'92*, K.P. Jantke (ed.), Lecture Notes in Computer Science, Springer Verlag, Berlin.

D.M. GABBAY (1985), 'Theoretical foundations for non-monotonic reasoning in expert systems'. In *Logics and Models of Concurrent Systems*, K.R. Apt (ed.), pp. 439-457, Springer Verlag, Berlin.

P. GÄRDENFORS (1988), *Knowledge in flux*, MIT Press, Cambridge, Massachusetts.

P. GÄRDENFORS (1990), 'Belief revision and nonmonotonic logic: two sides of the same coin?' In *Proc. Ninth European Conference on Artificial Intelligence*, pp. 768-773, Pitman, London.

E.M. GOLD (1967), 'Language identification in the limit', *Information and Control* **10**, pp. 447-474.

S. KRAUS, D. LEHMANN & M. MAGIDOR (1990), 'Nonmonotonic reasoning, preferential models and cumulative logics', *Artificial Intelligence* **44**, pp. 167-207.

T.M. MITCHELL (1982), 'Generalization as search', *Artificial Intelligence* **18**:2, pp. 203-226.

E.Y. SHAPIRO (1983), *Algorithmic program debugging*, MIT Press.

Y. SHOHAM (1987), 'A semantical approach to nonmonotonic logics'. In *Proc. Eleventh International Joint Conference on Artificial Intelligence*, pp. 1304-1310, Morgan Kaufmann, Los Altos, CA.

W. ZADROZNY (1991), *On rules of abduction*. IBM Research Report, August 1991.

OVERVIEW OF ITK RESEARCH REPORTS

No	Author	Title
1	H.C. Bunt	On-line Interpretation in Speech Understanding and Dialogue Sytems
2	P.A. Flach	Concept Learning from Examples Theoretical Foundations
3	O. De Troyer	RIDL*: A Tool for the Computer-Assisted Engineering of Large Databases in the Presence of Integrity Constraints
4	M. Kammler and E. Thijsse	Something you might want to know about "wanting to know"
5	H.C. Bunt	A Model-theoretic Approach to Multi-Database Knowledge Representation
6	E.J. v.d. Linden	Lambek theorem proving and feature unification
7	H.C. Bunt	DPSG and its use in sentence generation from meaning representations
8	R. Berndsen and H. Daniels	Qualitative Economics in Prolog
9	P.A. Flach	A simple concept learner and its implementation
10	P.A. Flach	Second-order inductive learning
11	E. Thijsse	Partical logic and modal logic: a systematic survey
12	F. Dols	The Representation of Definite Description
13	R.J. Beun	The recognition of Declarative Questions in Information Dialogues
14	H.C. Bunt	Language Understanding by Computer: Developments on the Theoretical Side
15	H.C. Bunt	DIT Dynamic Interpretation in Text and dialogue
16	R. Ahn and H.P. Kolb	Discourse Representation meets Constructive Mathematics

No	Author	Title
17	G. Minnen and E.J. v.d. Linden	Algorithmen for generation in lambek theorem proving
18	H.C. Bunt	DPSG and its use in parsing
19	H.P. Kolb	Levels and Empty? Categories in a Principles and Parameters Approach to Parsing
20	H.C. Bunt	Modular Incremental Modelling Belief and Intention
21	F. Dols	Compositional Dialogue Referents in Prase Structure Grammar
22	F. Dols	Pragmatics of Postdeterminers, Non-restrictive Modifiers and WH-phrases
23	P.A. Flach	Inductive characterisation of database relations
24	E. Thijsse	Definability in partial logic: the propositional part
25	H. Weigand	Modelling Documents
26	O. De Troyer	Object Oriented methods in data engineering
27	O. De Troyer	The O-O Binary Relationship Model
28	E. Thijsse	On total awareness logics
29	E. Aarts	Recognition for Acyclic Context Sensitive Grammars is NP-complete
30	P.A. Flach	The role of explanations in inductive learning
31	W. Daelemans, K. De Smedt and J. de Graaf	Default inheritance in an object-oriented representation of linguistic categories
32	E. Bertino and H. Weigand	An Approach to Authorization Modeling in Object-Oriented Database Systems
33	D.M.W. Powers	Multi-Modal Modelling with Multi-Module Mechanisms: Autonomy in a Computational Model of Language

No	Author	Title
34	R. Muskens	Anaphora and the Logic of Change*
35	R. Muskens	Tense and the Logic of Change
36	E.J. v.d. Linden	Incremental Processing and the Hierarchical Lexicon
37	E.J. v.d. Linden	Idioms, non-literal language and knowledge representation 1
38	W. Daelemans and A. v.d. Bosch	Generalization Performance of Backpropagation Learning on a Syllabification Task
39	H. Paijmans	Comparing IR-Systems: CLARIT and TOPIC
40	R. Muskens	Logical Omniscience and Classical Logic
41	P. Flach	A model of induction

ITK: P.O. BOX 90153 5000 LE TILBURG THE NETHERLANDS