Inductive, Abductive and Pragmatic Reasoning

John Bell
Applied Logic Group
Computer Science Department
Queen Mary and Westfield College
University of London
London E1 4NS
jb@dcs.qmw.ac.uk

Abstract

This paper gives a modern version of Pierce's distinction between induction and abduction, according to which they are both forms of pragmatic (or context-dependent) reasoning. A detailed example of induction and abduction in the formalisation and implementation of predictive and explanatory commonsense causal reasoning is then given.

1 Induction, Abduction, and Pragmatic Reasoning

Pierce distinguished between deductive, inductive and abductive forms of reasoning in the context of Aritstotle's syllogistic. Roughly, deduction consists in inferring a valid conclusion given the premisses, induction consists in inferring a reasonable (though not valid) conclusion given the premisses, and abduction consists in inferring an appropriate premise such that the conclusion is a valid consequence of it and the given premise. This definition has the merit of making clear what is given and what is inferred in each case. It also draws the distinction between deductive reasoning and the non-deductive reasoning involved in induction and abduction. However the definition is restricted to arguments of syllogistic form.

It is common in Philosophy, to use the term 'inductive argument' more liberally. For example, Swinburne [15] proposes the following definition: "An inductive argument is an argument which is not deductively valid but one in which, it is claimed, the premisses 'make it reasonable' for us to accept the conclusion." [15], p. 2. A special case of this is referred to as *enumerative induction* in philosophy and *rule induction* in AI; where a general rule in induced on the basis of a set of particulars. Much of our everyday commonsense reasoning is inductive; for example, we reason that the sun will rise tomorrow because it has always done so in the past and there is no reason to think otherwise. Inductive reasoning also plays a central part in more formal scientific reasoning; for example, in much of the reasoning which is intended to justify particular natural laws and in the predictions which are based on them. A correct inductive argument, according to Swinburne, is one in which the "premisses ... bear a favourable evidential relation to the conclusion", the "premisses make it reasonable for us to accept the conclusion in the sense, very roughly, that they report evidence favouring the truth of the conclusion, making it rational to believe the conclusion, making it likely or adding to the likelihood of its truth", p. 3. He sharpens this to "an argument ... [is] a correct inductive argument if its premisses make its conclusion more likely to be true that any equally detailed rival", p.4. Swinburne also notes that 'the argument will only yield probable knowledge if we know nothing else which affects the probability of the conclusion, apart from what is stated by the premisses ... the addition of new knowledge to the old knowledge may mean that I cannot make the same inductive inferences from the old knowledge. Hence I can only use an inductive argument to give me probable knowledge or reasonable belief, if the premisses contain all of my relevant knowledge", pp. 6-7. For futher discussion see [5; 15].

The term 'abduction' will also be interpreted in a wide sense. Following Harman [9], abduction can be characterised as inference to the best explanation. The givens are a set of premisses and a conclusion which is not a deductive consequence of the premisses, and the abductive task is to find the appropriate set of additional premisses such that the conclusion is a deductive consequence of the augmented premise set.

It is possible to use the idea of pragmatic entailment [2; 3] to give a general model-theoretic account of induction and abduction. Let $\mathcal{L}$ be a language and I be the class of all models of $\mathcal{L}$. A semantic meaning function is a function $[\![]\!]$ from $\mathcal{L}$ to $\wp I$. Intuitively, for sentence $\phi \in \mathcal{L}$, $[\![\phi]\!]$ represents the semantic, or literal, or context-independent meaning of ϕ. Then for each $\Theta \subseteq \mathcal{L}$, $[\![\Theta]\!] = \bigcap\{[\![\phi]\!] | \phi \in \Theta\}$. The semantic consequence relation induced by the (semantic) meaning function $[\![]\!]$ can then be defined in the usual way. For any $\Theta \subseteq \mathcal{L}$ and $\phi \in \mathcal{L}$, $\Theta \models \phi$ ("Θ (semantically) entails ϕ") iff $[\![\Theta]\!] \subseteq [\![\phi]\!]$. So semantic consequence is just the standard Tarkian notion. The analysis is then extended to pragmatic reasoning. A pragmatic meaning function is a function $[]$ from $\wp\mathcal{L}$ to $\wp I$; if ϕ is a sentence of $\mathcal{L}$ then $[\{\phi\}]$ is conveniently abbreviated to $[\phi]$. Intuitively for each $\Theta \subseteq \mathcal{L}$, $[\Theta]$ represents the pragmatic, or intended, or context-dependent meaning of Θ. Then $[]$ and $[\![]\!]$ induce a pragmatic consequence relation: for any $\Theta \subseteq \mathcal{L}$ and $\phi \in \mathcal{L}$, $\Theta \mid\!\approx \phi$ ("Θ pragmatically entails ϕ") iff $[\Theta] \subseteq [\![\phi]\!]$. Thus Θ pragmatically entails ϕ iff ϕ follows from the intended interpretation of Θ. Assuming that $\mathcal{L}$ contains the symbols $\top$, $\bot$ and $\wedge$, and that $[\![\top]\!] = I$, $[\![\bot]\!] = \emptyset$ and $[\![\phi \wedge \psi]\!] = [\![\phi]\!] \cap [\![\psi]\!]$, the properties of pragmatic entailment are follows:

(Normality) $\Theta \mid\!\approx \top$,
(Right Weakening)
If $\Theta \mid\!\approx \phi$ and $\phi \models \psi$ then $\Theta \mid\!\approx \psi$,
(And)
If $\Theta \mid\!\approx \phi$ and $\Theta \mid\!\approx \psi$ then $\Theta \mid\!\approx \phi \wedge \psi$,
(Consistency Preservation)
$\Theta \mid\!\approx \bot$ only if $\Theta \models \bot$.

In the case of induction, we wish to capture the idea that the premisses Θ represent a partial epistemic context and that induction is the process of extending Θ in an appropriate way and then reasoning semantically. Thus Θ induces ϕ if there is a reasonable expansion $\Theta \cup \Theta'$ of Θ which (semantically) entails ϕ. Intuitively Θ' contains premisses to the effect that Θ contains all of the relevant information, and premisses reflecting the expectations which arise given Θ. In order to formalise this, we require that the pragmatic function is veridical (for each $\Theta \subseteq \mathcal{L}$, $[\Theta] \subseteq [\![\Theta]\!]$) giving the additional property:

$$\text{(Identity)} \quad \Theta \mid\!\approx \Theta.$$

Thus, for consistent Θ $\Theta \mid\!\approx \phi$ if $\Theta \models \phi$ or there is some consistent extension $\Theta \cup \Theta'$ of Θ such that $\Theta \cup \Theta' \models \phi$. What counts as a reasonable extension can be further constrained by imposing additional conditions on the pragmatic function, with corresponding rationality postulates on the inductive consequence relation. Postulates of this kind are considered by Flach [7]. In many cases, the postulates of the system **P** of Kraus, Lehmann and Magidor [12] together with Consistency Presevation are appropriate; as for example is the case for the prediction theories discussed below. In other cases, weaker postulates may be appropriate; for example when the inductive arguments are expressed in Default Logic [13]; see [2] for a discussion. Deciding on an appropriate set of postulates represents the logical problem of induction. Picking a particular consequence relation which satisfies the postulates—picking a particular pragmatic function which satisfies the corresponding conditions—represents the pragmatic problem of induction.

In the case of abduction, we wish to capture the idea that the premisses Θ represent a partial epistemic context on the basis of which we are required to explain ϕ, and that abduction is the process of extending Θ in an appropriate way in order to do so. Abduction can thus be thought of as the process demonstrating that Θ induces ϕ by finding a particular Θ' such that $\Theta \cup \Theta'$ is consistent and $\Theta \cup \Theta' \models \phi$; where, intuitively, Θ' represents the explanation (or explanations) for ϕ given Θ. As with induction, the postulates of the system **P** together with Consistency Presevation are appropriate for many forms of abduction.

Thus both induction and abduction can be seen as forms of pragmatic, or context-dependent, reasoning. What distinguishes them is what is given and what has to be inferred.

2 Prediction is Induction, but Explanation is Abduction

In this section we illustrate the above discussion by outlining the approach to formalising commonsense causal reasoning developed in [6].

The Temporal Calculus, $\mathcal{TC}$, is based on Kleene's strong three-valued logic [11] and the temporal nonmonotonic extension of it in [4].

Kleene introduced the third truth value *undefined* in order to be able to describe computations which do not, in principle, terminate. Instead we adopt an epistemic resource-bounded interpretation. We are resource-bounded epistemically-limited reasoners, who can only know a tiny fragment of what there is to know. However, what matters in our everyday commonsense reasoning is that we know those facts which are relevant to us. Consequently given our limitations we concentrate our efforts on attempting to establish the truth or falsity of those facts that we consider relevant and ignore (treat as undefined) the truth values of the rest. However, given our limitiations, we do not always succeed in our attempts. On this interpretation then a sentence ϕ can be true (established by the reasoner(s) as being true), false (established by the reasoner(s) as being false) or undefined (ignored by the reasoner(s) as irrelevant, or unestablishable within the reasoner(s) resource limitations).

The syntax of Kleene's logic is identical with that of classical Predicate Calculus. The truth conditions of the logical constants yield a Boolean truth value wherever possible. Thus, for example, a sentence of the form $\phi \vee \psi$ is true if at least one of the disjuncts is true (even if the other is undefined).

As it stands, Kleene's logic is appropriate for representing monotonic reasoning with partial information. In order to be able to represent nonmonotonic reasoning, an *undefined* operator is added. Informally $?\phi$ states that the truth value of ϕ is undefined; that is, that neither the truth nor the falsity of ϕ is established. Then *weak affirmation*, *weak negation*, and *defined* operators are introduced:

$$\circ\phi \quad \stackrel{\text{def}}{=} ?\phi \vee \phi \qquad \text{“}\phi \text{ is not false”}$$

$$\bullet\phi \quad \stackrel{\text{def}}{=} ?\phi \vee \neg\phi \qquad \text{“}\phi \text{ is not true”}$$

$$!\phi \quad \stackrel{\text{def}}{=} \phi \vee \neg\phi \qquad \text{“}\phi \text{ is not undefined”}$$

Meta-level distinctions such as "ϕ is not false" and "ϕ is not true", and consequently non-monotonic inference rules, can be expressed in the object language; see, for example, the rules (1) and (2) given below.

In keeping with the resource-bounded interpretation, a new conditional operator is introduced. A sentence of the form $\phi \rightarrow \psi$ should be thought of as expressing a constraint which must be met if ϕ is true, but which can otherwise be ignored.

In order to represent time, an explicit temporal index is added to all atomic formulas of the underlying language (to all *Kleene atoms*); for example $On(B, L)(4)$ states that block B is on location L at time point 4. The earlier-than operator, $<$, allows comparison on time points, and quantification over time points and intervals is introduced.

The difference in aspect between facts and events is signified by means of the distinguished predicate *Occurs*. Thus a sentence of the form $\mathit{Occurs}(e)(t)$ states that a token of a primitive event of type e occurs at time t. Note that more than one (distinct) event (token) may occur simultaneously. Compound events can then be introduced by definition.

In order to gain comparable expressivity with the Situation Calculus, Kleene atoms and their negations, *Kleene literals*, are encoded as terms. For simplicity, autonomous naming is used, as there is no danger of confusion. In addition the distinguished predicates *True*, *Qual*, and *Aff* are introduced. The truth predicate, *True*, functions similarly to the $Holds$ relation of the Situation Calculus. Thus a sentence of the form $\mathit{True}(\lambda)(t)$ states that the Kleene literal (named by) λ is true at time t. As might be expected of a truth predicate, we have $\mathit{True}(\lambda)(t) \equiv \lambda(t)$ for any Kleene literal λ. A sentence of the form $\mathit{Qual}(e)(t)$ states that the event e is qualified at time t; that is, that there is reason to doubt that it will succeed, given that it occurs at t. A sentence of the form $\mathit{Aff}(\lambda)(t)$ states that the Kleene literal λ is affected at t; that is, that there is reason to doubt that its truth value persists beyond t.

Prediction Theories (PTs) are expressed in $\mathcal{TC}$. PTs may contain *boundary conditions*, *prediction rules*, *domain rules* and *default rules*.

The boundary conditions are unconditional sentences which state facts and the occurrence of events.

Prediction rules are temporally forwards-directed conditionals (conditionals in which, in any instance, the latest time point referred to in the antecedent temporally preceeds that referred to in the consequent) whose antecedents are typically hypothetical. These are used to define change. For example:

$$\forall e, t(\mathit{Pre}(e)(t) \wedge \mathit{Occurs}(e)(t) \wedge \bullet \mathit{Qual}(e)(t) \rightarrow \mathit{Post}(e)(t+1)) \tag{1}$$

$$\forall \lambda, t(\mathit{True}(\lambda)(t) \wedge \bullet \mathit{Aff}(\lambda)(t) \rightarrow \mathit{True}(\lambda)(t+1)) \tag{2}$$

Thus (1) is a general rule of change which states that if event e occurs at t and the necessary preconditions of e are true at t and e is not qualified at t, then infer that the postconditions of e are true at $t+1$; that is, that e succeeds at t. Rule (2) is a general rule of inertia, which states that if the Kleene literal (named by the term) λ is true at t and (the Kleene literal named by) λ is not affected at t then infer that (the Kleene literal named by) λ is true at $t+1$. Given that, for any Kleene literal λ, $\mathit{True}(\lambda)(t)$ is equivalent to $\lambda(t)$, (2) is equivalent to the following higher-order sentence:

$$\forall \lambda, t(\lambda(t) \wedge \bullet \mathit{Aff}(\lambda)(t) \rightarrow \lambda(t+1)) \quad (3)$$

In practice it is convenient to use (3) as an abbreviation for (2).

Domain rules are categorical temporally undirected (bi-)conditionals which are used for a variety of purposes, including the definition of domain constraints and ramifications.

Finally, default rules are temporally undirected conditionals with hypothetical antecedents which are used to do default reasoning at a time point(or within an interval. The intention is to allow a limited amount of default reasoning, not to incorporate Default Logic; it is thus assumed that these rules can be stratified.

The pragmatics of PTs are given by *prioritised chronological minimisation*. This captures three necessary properties of predictive causal reasoning: it is rule-governed (albeit by defeasible rules), it is temporally forwards-directed, and potential change takes precedence over inertia; that is, if an action can succeed, it does succeed. Prioritised chronological minimisation is realised by exploiting the partiality of the Kleene basis and chronologically minimising *information* (what is defined/established). In order to give preference to potential change, qualifications (reasons why an event may not occur successfully) are minimised in preference to affectations (reasons why a fact should not persist) at each time point. Informally, two models M and M' are said to be comparable if they differ at most on the interpretation of ordinary literals or on the interpretation of *Occurs*, *Qual* or *Aff* literals. Then $M \prec_{pcld} M'$ (M is preferrred to M') if M and M' are comparable and there is a time point t such that M and M' agree up to t, and:

- at least one more ordinary or *Occurs* literal is established in M' at t, or
- M and M' agree on ordinary and *Occurs* literals, and at least one more *Qual* literal is established in M' at t, or
- M and M' agree on all ordinary, *Occurs*, and *Qual* literals, and at least one more *Aff* literal is established in M' at t.

Definition 2.1 *A model M is a* prioritised chronologically least defined *model (*p.c.l.d. model*) of ϕ if $M \models \phi$ and there is no other model M' such that $M' \models \phi$ and $M' \prec_{pcld} M$. Likewise, M is a p.c.l.d. model of a set of sentences Θ if $M \models \Theta$ and there is no other model M' such that $M' \models \Theta$ and $M' \prec_{pcld} M$.*

Definition 2.2 *A PT Θ* predicts *a sentence ϕ, written $\Theta \approx_{pcld} \phi$, if ϕ is true in all p.c.l.d. models of Θ.*

In order to illustrate inductive and abductive reasoning in this framework we discuss prediction and explanation and give a simple example of each. In the sequel we adopt the convention that free variables are implicitly universally quantified.

The frame (or inertia, or persistence) problem is the problem of efficiently representing what *does not* change as result of an event. In other words, we want to represent the commonsense inference that facts persist, other things being equal; where other things are equal if no event occurs which affects those facts. In order to do so, we use a general inertia rule which is analogous to those used in Situation Calculus-based approaches. For any Kleene literal λ:

$$\lambda(t) \wedge \bullet \mathit{Aff}(\lambda)(t) \rightarrow \lambda(t+1) \quad (4)$$

In simple cases it is sufficient to use the inertia rule to conclude "by default" that facts persist. Particular applications of the rule are implicitly blocked by "back propagation".

Example 2.3 The Yale Shooting Problem [8]. At time 1 a gun is loaded and pointed at Fred. Nothing relevant happens at time 2. At time 3 the gun is fired. Is Fred dead as a result? Intuitively we think so because we infer that the gun remains loaded until it is fired. The problem is to represent this commonsense inference formally.

Let Θ be the PT $\{(4),(5),(6)\}$, where:

$$\mathit{Alive}(1) \wedge \mathit{Loaded}(1) \wedge \mathit{Occurs}(\mathit{Shoot})(3) \quad (5)$$

$$\mathit{Loaded}(t) \wedge \mathit{Occurs}(\mathit{Shoot})(t) \rightarrow \neg \mathit{Alive}(t+1) \quad (6)$$

Then it is easy to see that $\Theta \approx_{pcld} Loaded(3) \wedge Alive(3) \wedge \neg Alive(4)$. Let M be a p.c.l.d. model of Θ. Then $M \models ?\mathit{Aff}(L)(1)$ and $M \models ?\mathit{Aff}(A)(1)$, so by (4), $M \models Loaded(2)$ and $M \models Alive(2)$. Similarly $M \models Loaded(3)$ and $M \models Alive(3)$. Rule (6) now gives $M \models \neg Alive(4)$. Rule (6) also gives $M \models \mathit{Aff}(A)(3)$ thus blocking the application of rule (4) at time 3 in which λ is A. □

As suggested by the example, the (finite restrictions of) the intended model(s) of appropriate PTs can be generated by starting with an initial model, $M/0$, and then using the rules as detailed instructions for extending each M/t to $M/t+1$. Initial work on this model-building implementation is discussed in [3; 4].

PTs can also be used as a basis for the representation of explanatory commonsense causal reasoning. Standardly [10; 14], given a theory Θ such that $\Theta \not\models \phi$, abduction is presented as the process of finding appropriate minimal sets Θ_i such that each $\Theta \cup \Theta_i$ is consistent and $\Theta \cup \Theta_i \models \phi$. The various Θ_i are then the explanations for ϕ given Θ. We change this definition slightly and seek appropriate minimal Θ_i such that $\Theta \cup \Theta_i \approx_{pcld} \phi$. Using $\approx_{pcld}$ rather than $\models$ as the consequence relation should result in fewer, more relevant, explanations (as the obvious nondeductive conclusions are already included in the $\approx_{pcld}$ relation). Furthermore, the present approach suggests that each Θ_i can be *generated* by starting at the initial time point and then reasoning forwards in time assuming only what is necessary to predict the conclusion at each interim time point (and then checking that the conclusion is true at the final time point). In order to count as reasonable explanations, the assumptions, or *abducibles*, should be *relevant* and *non-trivial*. For present purposes, relevant assumptions can be taken to be instantiations of literals occurring in the antecedents of rules of the theory Θ. While nontriviality can be taken to be the requirement that $\Theta_i \not\models \phi$. Explanation is thus viewed as *augmented prediction*.

Example 2.4 The Stanford Murder Mystery [1]. In this variant of the Yale Shooting Problem, Fred is alive at time 1. As before, if the gun is loaded when it is fired, then Fred is dead as a result. We are also told that a shoot event occurs at time 1 and that Fred is dead at time 3. When and why did Fred die? It seems that there are two possible explanations. One is that the gun was already loaded at time 1 and Fred was dead at time 2 as a result of being shot at time 1. The other, less likely but still plausible, is that the gun was loaded at time 1 and fired at time 2.

Let Θ be the PT $\{(4), (7), (8), (9), (10)\}$; where:

$$Alive(1) \wedge Occurs(Shoot)(1) \wedge \neg Alive(3) \quad (7)$$

$$1 \leq t \leq 2 \rightarrow !\mathit{Aff}(Alive)(t) \quad (8)$$

$$Occurs(Load)(t) \rightarrow Loaded(t+1) \quad (9)$$

$$Loaded(t) \wedge Occurs(Shoot)(t) \rightarrow \neg Alive(t+1) \quad (10)$$

The abductive task is to find the minimal Θ_i such that $\Theta \cup \Theta_i \approx_{pcld} \neg Alive(3)$. As suggested, this can be done by reasoning forwards in time from the initial time point, and assuming only what is necessary to establish the conclusion at each time point. Thus at time 1 we have $\Theta_1 = \{Loaded(1)\}$, and at time 2 we have $\Theta_2 = \{Occurs(Load)(1), Occurs(Shoot)(2)\}$ as possible explanations. The gun was loaded at time 1 and so, as a result of the shoot event, Fred was dead at time 2. Alternatively, Fred was alive at time 2 and died as a result of a load event at time 1 and a shoot event at time 2.

In Baker's original formulation of the problem, a wait event occurs at time 2. If we add a rule to this effect:

$$\neg \exists e\, Occurs(e)(2), \quad (11)$$

then we are left with Θ_1 as the only explanation. □

Acknowledgements

I would like to thank the participants of the ECAI-96 Workshop on *Abductive and Inductive Reasoning* for helpful discussions and comments.

References

[1] A. Baker (1991) Nonmonotonic Reasoning in the Framework of the Situation Calculus. *Artificial Intelligence* 49 pp. 5-23.

[2] J. Bell. Pragmatic Logics. Proceedings of KR '91, Morgan Kaufmann, 1991, pp. 50-60. Revised and extended in 1992.

[3] J. Bell. Pragmatic Reasoning; a Model-Based Theory. In M. Masuch and L. Polos (eds.), *Applied Logic, How, What and*

Why? A selection of papers from the Applied Logic Conference, Amsterdam, December 1992. Kluwer Academic Publishers, Amsterdam, 1995, pp. 1-28.

[4] J. Bell. A Model-Based Approach to Predictive Causal Reasoning. In: P. Doherty (ed.), *Partiality, Modality, and Non-Monotonicity*. FoLLI Studies in Logic Language and Information series, CSLI Publications, Stanford, 1996, pp. 169-195.

[5] J. Bell. Induction in Philosophy and AI. Working Notes of the ECAI-96 Workshop on *Abductive and Inductive Reasoning*, M. Denecker, L. De Raedt, P. Flach and T. Kakas (eds.), pp. 7-9.

[6] J. Bell. Prediction Theories. Manuscript, 1997.

[7] P. Flach. Rationality Postulates for Induction. Proceedings of the Sixth Conference on Rationality and Knowledge, TARK'96, Y. Shoham (ed.), pp. 267-281.

[8] S. Hanks and D. McDermott (1986) Default Reasoning, Nonmonotonic Logics, and the Frame Problem. Proceedings of AAAI'86, pp. 328-333.

[9] G. Harman. Inference to the best explanation. *Philosophical Review* 74, pp. 88-95.

[10] A.C. Kakas, R.C. Kowalski and F. Toni. Abductive Logic Programming. *Journal of Computational Logic* 2(6), 1992, pp. 719-770.

[11] S.C. Kleene. *Introduction to Metamathematics*. North-Holland, Amsterdam, 1952.

[12] S. Kraus, D. Lehmann and M. Magidor. Nonmonotonic Reasoning, preferential models and cumulative logics. *Artificial Intelligence* 44 1990, pp. 167-207.

[13] R. Reiter. A logic for default reasoning. *Artificial Intelligence* 13 1980, pp. 81-132.

[14] M. Shanahan. Explanation in the Situation Calculus. Proceedings of IJCAI'93, pp. 160-165.

[15] R. Swinburne (ed.) *The Justification of Induction*. Oxford Readings in Philosophy series, Oxford University Press, Oxford, 1974.