



THEORIA. Revista de Teoría, Historia y
Fundamentos de la Ciencia

ISSN: 0495-4548

theoria@ehu.es

Universidad del País Vasco/Euskal Herriko
Unibertsitatea
España

NIINILUOTO, Ilkka

Structural Rules for Abduction

THEORIA. Revista de Teoría, Historia y Fundamentos de la Ciencia, vol. 22, núm. 3, 2007, pp. 325-329

Universidad del País Vasco/Euskal Herriko Unibertsitatea
Donostia-San Sebastián, España

Available in: <http://www.redalyc.org/articulo.oa?id=339730804008>

- How to cite
- Complete issue
- More information about this article
- Journal's homepage in redalyc.org

redalyc.org

Scientific Information System
Network of Scientific Journals from Latin America, the Caribbean, Spain and Portugal
Non-profit academic project, developed under the open access initiative

Structural Rules for Abduction

Ilkka NIINILUOTO

BIBLID [0495-4548 (2007) 22: 60; pp. 325-329]

ABSTRACT: Atocha Aliseda's *Abductive Reasoning* (2006) gives a structural characterization of the "forward" explanatory reasoning from a theory to observational data. This paper asks whether there are any interesting structural rules for the "backward" abductive reasoning from observations to explanatory theories. Ignoring statistical cases, a partial explication of abduction is converse deductive explanation: b is abducible from e iff b deductively explains e . This relation of abducibility trivially satisfies Converse Entailment (if b entails e , then b is abducible from e), but it does not generally satisfy Converse Consequence (if b is abducible from e and g entails b , then g is abducible from e), since deductive explanation is not always transitive.

Keywords: abduction, deduction, explanation, induction, non-monotonic logic, structural rules.

One of the important tools used in Atocha Aliseda's *Abductive Reasoning* (2006) is the analysis of logical inference by means of structural rules. Such rules express formal properties of various kinds of reasoning by general principles of inference. Aliseda gives a powerful structural characterization of the "forward" explanatory reasoning from a theory to observational data. In this note, I ask whether there are any interesting structural rules for the "backward" abductive reasoning from observations to explanatory theories.

Deductive Explanation

Aliseda's Chapter 3 gives a structural characterization of "consistent explanatory abduction" $\Theta \mid \alpha \Rightarrow \varphi$, defined by the conditions (i) $\Theta, \alpha \vdash \varphi$, and (ii) Θ, α are consistent. Here Θ is a theory, α an initial condition (the explanans), and φ the explanandum. In Chapter 5, Aliseda discusses Hempel's DN-model of scientific explanation as a form of consistent explanatory abduction $t \mid c \Rightarrow e$, where it is also required that (iii) not $t \vdash e$, and (iv) not $c \vdash e$. *Deductive explanation* differs from mere deduction, since it includes additional conditions which exclude trivial cases of inconsistencies, ad hoc explanation and self-explanation. Hence, the relation \Rightarrow fails to satisfy conditions of monotonicity and cut, which are typical of deduction, but restricted or modified forms of these principles may be valid.

Aliseda's HD Monotonicity allows consistent strengthening of an explanatory theory:

$$(1) \quad \frac{t \mid a \Rightarrow e \quad t, b \mid d \Rightarrow e \quad t, b, a \text{ is consistent}}{t, b \mid a \Rightarrow e}$$

The HD Cut principle



$$(2) \quad \frac{t \mid a \Rightarrow b \quad t \mid b \Rightarrow e \quad \text{not } a \quad e}{t \mid a \Rightarrow e}$$

states that explanation is almost transitive. Other valid principles include Disjunction of Theories

$$(3) \quad \frac{t_1 \mid c \Rightarrow e \quad t_2 \mid c \Rightarrow e}{(t_1 \vee t_2) \mid c \Rightarrow e}$$

and Conjunction of two Explananda

$$(4) \quad \frac{t \mid c \Rightarrow e_1 \quad t \mid c \Rightarrow e_2}{t \mid c \Rightarrow (e_1 \wedge e_2)}$$

But if explanation is required to be minimal (i.e., without irrelevant additional parts) or preferred (i.e., the best among rival explanations), such structural principles cannot be defended.

To simplify notation, let tEe mean that “ t deductively explains e ”, where the explanans t may be a conjunction of theoretical assumptions and initial conditions. Conditions (iii) and (iv) are special cases of Raimo Tuomela’s (1973) non-comparability requirement: in deductive explanation the explanandum should be logically independent of the ultimate truth-functional conjunctive components of the explanans. Let us denote this condition of non-comparability by $\text{inc}(t, e)$. Let $\text{cons}(t, e)$ mean that $t \wedge e$ is consistent. Then deductive explanation can be defined by

$$(E) \quad tEe \text{ =}_{\text{df}} (t \mid e) \ \& \ \text{cons}(t) \ \& \ \text{inc}(t, e).$$

Tuomela’s non-comparability condition makes explanation irreflexive (cf. Aliseda, 2006, p. 138), so that even Aliseda’s Conditional Reflexivity (*op. cit.*, p. 76) fails:

$$(5) \quad \text{Irreflexivity: not } tEt.$$

Irreflexivity entails that deductive explanation is asymmetric, since eEt and tEe would imply that e and t are logically equivalent:

$$(6) \quad \text{Asymmetry: If } tEe, \text{ then not } eEt.$$

The principle (1) can now be rewritten in the form:

$$(1') \quad \text{Monotonicity: If } tEe, \text{ inc}((t \wedge b), e) \text{ and } \text{cons}(t, b), \text{ then } (t \wedge b)Ee$$

This implies

$$(7) \quad \text{Conjunction of Explanantia: If } t_1Ee, t_2Ee, \text{ and } \text{cons}(t_1, t_2), \text{ then } (t_1 \wedge t_2)Ee.$$

Conditions (2)-(4) can now be rewritten:

$$(2') \quad \text{Transitivity: If } uEt \text{ and } tEe, \text{ and } \text{inc}(u, e), \text{ then } uEe$$

- (3') *Disjunction*: If $t_1 E e$ and $t_2 E e$, then $(t_1 \vee t_2) E e$
- (4') *Conjunction of Explananda*: If $t E e_1$ and $t E e_2$, then $t E (e_1 \wedge e_2)$.

Inducibility

In my early work, I used a similar approach in the study of inductive reasoning, where the relation eIb means that “ b is inducible from e ” (either in the sense of *inductive confirmation* or *inductive acceptance*) (see Niiniluoto 1972; Niiniluoto and Tuomela 1973). Even though I had not yet heard about non-monotonic logic, I argued against the monotonicity and transitivity of induction:

- (8) *Monotonicity*: If eIb , then $(e \wedge b)Ib$
- (9) *Transitivity*: If eIb and bIg , then eIg .

It was shown already in 1945 by Carl G. Hempel that no non-trivial notion of qualitative confirmation can satisfy at the same time the principles

- (En) *Entailment*: If $e \supset b$, then eIb
- (CC) *Converse Consequence*: If eIb and $g \supset b$, then eIg .

(See Hempel, 1965.) Similarly, the following principles exclude each other:

- (CE) *Converse Entailment*: If $b \supset e$, then eIb
- (SC) *Special Consequence*: If eIb and $b \supset g$, then eIg .

It is well-known that the *high probability criterion* of confirmation (eIb iff $P(b/e) > q \geq 1/2$) satisfies Entailment En and Special Consequence SC. Howard Smokler (1968) suggests that these conditions are characteristic to “enumerative” and “eliminative” reasoning. Aliseda’s proposal of combining inductive and deductive explanation, or “deductive cut on the explanans” (*op. cit.*, p. 145), and its justification by appeal to high probability, is an instance of SC.

On the other hand, the positive relevance criterion of confirmation (eIb iff $P(b/e) > P(b)$) satisfies Converse Entailment CE, but not conditions En, SC, and CC. Therefore, it does not completely fit Smokler’s idea that there is a form of “abductive inference” which satisfies CE and CC. Modifying an idea of Brody, it was suggested in Niiniluoto and Tuomela (1973) that one could replace the relation of deducibility by the stronger relation of deductive explanation E in CE and CC:

- (CE*) If $b E e$, then eIb .
- (CC*) If eIb and $g E b$, then eIg .

Then these two conditions might define Smokler’s “abductive inference”. Moreover, if En* and SC* are modified in the same way, then Hempel’s arguments do not any more count against combining En* and CC* or CE* and SC*. (For example, $t \& u$ does not explain t , and t does not explain $t \vee u$.)

Abductive Inference

Charles S. Peirce characterized abduction as an inference of a cause from its effect. Its conclusion is a hypothesis which is adopted “for the sake of explanation”. Peirce’s notion of explanation covers both deductive and inductive-statistical arguments (see Niiniluoto 1999). Thus, abductive or retroductive reasoning proceeds in a direction which is opposite to deductive or inductive explanatory arguments from causes to effects. Peirce’s best-known general formulation of abduction is the following:

The surprising fact C is observed.

(A) But if A were true, C would be a matter of course.

Hence, there is reason to suspect that A is true.

(Cf. Aliseda, *op. cit.*, p. 36.) In the weak sense discussed here, abduction is thus *reasoning from a fact to one of its potential explanations*. The stronger sense, where inference is to the *minimal* explanation or the *best* explanation of a fact, is beyond the scope on my remarks. It is assumed that the triggering fact for abduction is novelty, rather than anomaly, in Aliseda’s sense (*op. cit.*, p. 47). I shall also ignore the reinterpretation of (A), where the conclusion states that theory A is truthlike (see Niiniluoto 2005).

Ignoring statistical cases, a simple partial explication of Peirce’s schema (A) would be *converse deduction*:

(CD) $eCDb =_{df} b \rightarrow e \text{ \& not } \neg e$.

This definition is known also as “deductive support” (Niiniluoto and Tuomela 1973, p. 221; Kuipers 2000). It has the virtue that it covers both cases of deductive explanation and deductive prediction. It is a trivial consequence that CD satisfies CE. By the transitivity of deduction \rightarrow , it satisfies also CC. Hence, CD is a candidate for Smokler’s notion of “abductive inference”.

Another partial explication, which is more faithful to Peirce’s (A), is the definition of abduction as *converse deductive explanation*: b is *abducible* from e iff b deductively explains e . In other words,

(A) $eAb =_{df} bEe$.

This definition has been proposed in Niiniluoto and Tuomela (1973), p. 227. Again, the relation A of abducibility trivially satisfies CE*:

(10) If bEe , then eAb .

Relation A does not generally satisfy CC*, since deductive explanation is not always transitive. But by (2’) we have a conditional transitivity property for A :

(11) If eAb and bAu , and $\text{inc}(u, e)$, then eAu .

This rule implies a conditional converse consequence principle for A :

(12) If eAb and uEb , and $\text{inc}(u, e)$, then eAu .

As E differs from \rightarrow , abducibility does not satisfy the original formulations of CE and CC. It is also clear that it does not satisfy E , E^* , SC , or SC^* .

Abductive inference is non-monotonic, since the following principle is not valid:

(13) If eAh , then $(e \wedge b)Ah$.

Namely, if b explains e , then it is not always the case that b explains $(e \wedge b)$, even when b is consistent with b . The latter task of explanation is clearly more demanding than the former. But by (4') the following conjunction principle is valid:

(14) If e_1Ah and e_2Ah , then $(e_1 \wedge e_2)Ah$.

Another conjunction principle, which is not generally valid for inductive support I, is problematic:

(15) If eAb_1 and eAb_2 , then $eA(b_1 \wedge b_2)$,

since the conjunction of two rival explanations may be inconsistent. By (7) we have instead of (15):

(16) If eAb_1 , eAb_2 , and $\text{cons}(b_1, b_2)$, then $eA(b_1 \wedge b_2)$.

Further, principle (3') implies that

(17) If eAb_1 and eAb_2 , then $eA(b_1 \vee b_2)$.

Most of the principles for A given here would fail, if the notion of abduction would also cover cases of converse inductive explanation. But this issue has to be left for another occasion. What we have done so far already illustrates the stimulus of Atocha Aliseda's fine monograph.

REFERENCES

- Aliseda, A. (2006). *Abductive Reasoning: Logical Investigations into Discovery and Explanations*. Dordrecht: Springer.
- Hempel, C.G. (1965). *Aspects of Scientific Explanation*. New York: The Free Press.
- Kuipers, T. (2000). *From Instrumentalism to Constructive Realism: On some Relations between Confirmation, Empirical Progress and Truth Approximation*. Dordrecht: Kluwer.
- Niiniluoto, I. (1972). "Inductive Systematization: Definition and a Critical Survey", *Synthese* 25, 25-81.
- (1999). "Defending Abduction", *Philosophy of Science (Proceedings)* 66, S436-S451.
- (2005). "Abduction and Truthlikeness", in R. Festa, A. Aliseda, and J. Peijnenburg (eds.), *Confirmation, Empirical Progress, and Truth Approximation*. Amsterdam: Rodopi, pp. 255-275.
- and R. Tuomela (1973). *Theoretical Concepts and Hypothetico-Inductive Inference*. Dordrecht: D. Reidel.
- Smokler, H. (1968). "Conflicting Conceptions of Confirmation", *Journal of Philosophy* 65, 300-312.
- Tuomela, R. (1973). *Theoretical Concepts*. Berlin: Springer-Verlag.

Ilkka NIINILUOTO has made his academic career at the University of Helsinki: Ph.D. in 1974, Associate Professor of Mathematics in 1973-1977, Professor of Theoretical Philosophy in 1977, and Rector in 2003. Since 1975 he has been the President of the Philosophical Society of Finland, and since 1998 member of the Steering Committee of FISP. Niiniluoto's main works *Is Science Progressive?* (1984), *Truthlikeness* (1987), and *Critical Scientific Realism* (1999) deal with philosophy of science, but his interests include also philosophical logic, epistemology, philosophy of technology, philosophy of culture, ethics, and social philosophy.

ADDRESS: Department of Philosophy, P.O. Box 9, 00014 University of Helsinki, Finland. E-mail: ilkka.niiniluoto@helsinki.fi. Web: <http://www.helsinki.fi/filosofia/filo/henk/niiniluoto.htm>.