May 22nd, 9:00 AM - May 25th, 5:00 PM

Commentary on: Jean H. M. Wagemans' "The assessment of argumentation based on abduction"

Thomas F. Gordon
Fraunhofer FOKUS

Commentary on: Jean Wagemans' "The assessment of argumentation based on abduction"

THOMAS F. GORDON

Fraunhofer FOKUS
Berlin
Germany
thomas.gordon@fokus.fraunhofer.de

1. RELATIONSHIP BETWEEN ABDUCTION AND DEDUCTION

Professor Wagemans claims that:

> Abduction is a type of reasoning in which the presence of a certain cause is inferred from the presence of its effects. The type of reasoning is conventionally valid in various institutionalized contexts, for example the practice of medical diagnosis and that of scientific explanation, but logically invalid because it affirms the consequent.

This definition combines two perspectives on abduction: 1) a method for producing and comparing explanatory hypotheses of observations and 2) a deductively invalid form of reason with many applications, including but not limited to explaining observations.

Although abduction in the first sense, as a method for producing and comparing explanations of hypotheses, is typically formalized using deductively invalid inference patterns, it is surely possible to formulate this kind of reasoning in a sound way, for example using the following scheme:

Minor. E has been observed.
Major. H causes E.
Conclusion. H explains E.

The conclusion here means only that H would be *an* explanation of E, not necessarily the *best*. Thus this scheme models what Wagemans calls the "generative" kind of abduction.

This scheme is intended to be strict, not defeasible. The only way to attack arguments using this scheme is to show that one or both of the premises are false.

An alternative to formulating the scheme as an inference rule would be to represent it as a formula in a first-order theory:

$$\forall\ E\ H\ .\ E \text{ has been observed} \land H \text{ causes } E \Rightarrow H \text{ explains } E$$

Either way, the trick here of course is to conclude only that H is *a* possible explanation of E, rather than that H is the *best* explanation, let alone *true*. This weak claim can be strict. It need not be retractable with further information, especially

Mohammed, D., & Lewiński, M. (Eds.). *Virtues of Argumentation. Proceedings of the 10th International Conference of the Ontario Society for the Study of Argumentation (OSSA), 22-26 May 2013.* Windsor, ON: OSSA, pp. 1-6.

when used in combination with some method for comparing explanations to find the best ones.

This scheme for abduction is close to Pierce's formulation, cited in Wagemans' paper:

> The surprising fact, E, is observed; But if H were true, E would be a matter of course,
> Hence, there is reason to suspect that H is true.

The schema variables have been renamed to coincide with my version. Notice that neither version of the scheme concludes with H being the best explanation or true.

Here's a reconstruction of Pierce's bean example using this deductive version of abduction:

Minor. White beans on the floor have been observed.
Major. The beans fell out of this bag of white beans causes white beans on the floor have been observed.
Conclusion. The beans fell out of this bag of white beans explains white beans on the floor have been observed.

We have illustrated how abduction in the sense of a method for generating hypotheses for observations can be formulated as deductive valid inference. Let us now give an example of an entirely formal model of abduction that is not limited in its applicability to explaining observations.

Perhaps one of the most influential formalizations of abduction in the field of Artificial Intelligence is David Poole's Theorist system (Poole, 1988), which was developed not to model the process of explaining observations but rather as a framework for nonmonotonic reasoning. Poole's thesis was that nonmonotonic reasoning does not require a different logic, but just a different way to use classical logic: "if one allows hypothetical reasoning then there is no need to define a new logic to handle nonmonotonic reasoning."

Poole defined abduction as follows. Let F be a set of closed formulas of first-order logic, representing the "facts" of a situation, and Δ be a set of "possible hypotheses". Abduction is defined as the process of deriving the maximal subsets D of instantiations of Δ such that F ∪ D is consistent in classical logic. The deductive closure of such a F ∪ Δ is called an *extension* of (F, Δ). There can be no extension as well as multiple extensions.

Poole reconstructed the standard Tweety example of the nonmonotonic logic community using this framework as follows:

$$\Delta = \{bird(x) \Rightarrow flies(x)\}$$
$$F = \{ \forall x . emu(x) \Rightarrow bird(x), \forall x . emu(x) \Rightarrow \neg flies(x), emu(Polly), bird(Tweety) \}$$

bird(Tweety) is in an extension of (F,Δ), but not flies(Polly), because including bird(Polly) ⇒ flies(Polly) in D would allow both flies(Polly) and ¬flies(Polly) to be derivable, causing an inconsistency.

Notice that the Tweety example is not an example of generating a hypothesis to explain an observation. Tweety has been observed, we can presume, to be a bird, but the flies (Tweety) is not derived to explain this observation, but rather as a defeasible consequence of being a bird. Poole's choice of this example makes clear that his model of abduction is not intended to be used only to generate explanations of observations. His model can however *also* be used for this purpose.

2. SIMILARITIES BETWEEN ABDUCTION AND PRACTICAL REASONING

Professor Wagemans notes that

> Pierce in a later work describes abduction as a key operation in the process of generating a hypothesis that functions as an explanation of certain facts.

and then goes on to describe a process model for finding the best explanations of observations, consisting of these three steps:

1. Gather hypotheses (explanations).
2. Rate the quality of the explanations.
3. Choose one of the best explanations.

Here I would only like to point out the apparent similarity between this procedural conception of abduction and practical reasoning. Starting with a goal, rather than an observation, a procedure for practical reasoning might look like this:

1. Gather, for example via brainstorming, alternative possible course of action for realizing the goal.
2. Rate the quality of the alternative courses of action.
3. Choose one of the best courses of action.

My question is: Is this similarity more then superficial? Might abduction and practical reasoning be specializations of a more general form of reasoning?

3. ON THE DEFEASIBILITY OF ARGUMENTATION SCHEMES

Professor Wagemanns writes:

> Argumentation theorists widely use the term ‘argument(ation) scheme’ in order to describe various types of defeasible arguments.

Some people, including Henry Prakken and myself, view argumentation schemes as a *generalization* of (deductive) inference rules. From this perspective, argumentation schemes are typically defeasible, but not necessarily so. All of the strict inference rules of a natural deduction calculus for classical logic, for example, may also be viewed as argumentation schemes.

4. ON THE TYPES OF CONCLUSIONS OF DIFFERENT FORMS OF INFERENCE

Professor Wagemans points out the common view that one difference between abduction and induction concerns the types of their conclusions:

> Other scholars emphasize the difference between abduction and induction. They note that these types of reasoning produce a different type of conclusion, since in the case of induction the conclusion is of a general nature, whereas in the case of abduction the conclusion is of a particular nature.

Here I would only like to draw attention to criticisms of this view. For example Walton, in his "Fundamentals of Critical Argumentation" (Walton, 2006, p. 67) writes:

> There is a common misconception that deductive argumentation is general to the specific, while inductive reasoning always goes from the specific to the general.

Walton claims the following is a counterexample, of induction with a specific conclusion:

Premise.	Most students who graduated from Bohemond College after 1995 took a course on critical thinking.
Premise.	Elaine was a student who graduated from Bohemond College after 1995.
Conclusion.	Elaine took a course on critical thinking.

The abductive conclusion of Poole's Tweety example, presented earlier in this comment, is the instantiation

$$\text{bird(Tweety)} \Rightarrow \text{flies(Tweety)}$$

of the hypothesis

$$\text{bird}(x) \Rightarrow \text{flies}(x)$$

in the set of hypotheses Δ.

It is unclear to me how this conclusion would be viewed by those who consider conclusions of abduction to be "of a particular nature". Would this formula be viewed as "particular", because it is ground, or "general", because it is a material implication?

5. DIFFERING CONCEPTIONS OF ARGUMENT

Not being as familiar as I probably should be with the pragma-dialectic conception of argument, I needed to read the paper a couple of times to understanding statements such as the following:

> Abduction defined as 'inference to the best explanation' involves a standpoint that can be formulated as "Hi is the best explanation of E" and an argument that can be formulated as "Of candidate explanations H1, …, Hn of E, Hi meets criteria C1, …, Cn best".

It was not immediately clear to me how the proposition "Of candidate explanations H1, …, Hn of E, Hi meets criteria C1, …, Cn best" can be understood to be an argument. In the field of computational models of argument, several conceptions of arguments are common:

- Argument as a single (defeasible) inference step, i.e. a pair consisting of a set of premises and a conclusion, where the premises and conclusion are propositions. Such arguments may (but need not be) instantiations of argumentation schemes of the kind Doug Walton has been developing (Walton 1996; Walton et al. 2008). Besnard and Hunter adopt a similar view of the structure of argument in their deductive conception of argument (Besnard and Hunter 2008), except that they require the conclusion of an argument to be a deductive consequence, in classical logic, of the premises of the argument.
- Argument as a (defeasible) proof, i.e. a chain or tree of inference steps. These are the kinds of arguments that are visualized using Beardsley/Freeman argument diagrams, for example in Walton's textbook "Fundamentals of Argument" (Walton 2006). This is also the conception of argument used by Henry Prakken in his ASPIC+ model of structured argument (Prakken 2010).

It seems to me that pragma-dialectics adopts yet another conception of argument, to mean a kind of minor premise of an inference step. Let's take another look at Wagemann's pattern of argumentation based on abduction in Section 4 of his paper:

1 It may be hypothesized that X1
1.1 It is observed that Y
    1.1' Of possible explanations X1 – Xn, X1 is the best explanation of Y
        1.1'.1 X1 meets criteria C1 – Cn with scores S1 – Sn
        1.1'.1' Decision rule R applies

If I understand correctly, when this pattern is applied, the resulting argument is an instantiation of "It is observed that Y" and the conclusion of the argument is an instantiation of "It may be hypothesized that X1". This pattern could be reconstructed as a Walton-style argumentation scheme as follows:

Minor.	It is observed that Y
Major.	Of possible explanations X1 – Xn, X1 is the best explanation of Y.
Conclusion.	It may be hypothesized that X1.

If I understand correctly, in pragma-dialectics the major premise (in this reconstruction) is called the "justificatory force" of the argument. The reasons for preferring this terminology in pragma-dialectics are not clear to me, but I suspect they are based on the observation that in natural language arguments typically only the minor premise is explicitly stated. The major premise is often left implicit, because is it assumed to be common knowledge already accepted by the audience. That is, the argument is an enthymeme. For example, the classical example of a syllogism would be expressed enthymematically as "Socrates is mortal since he is human", not as "Socrates is mortal because all men are mortal".

The remainder of the pattern can be handled by a second argumentation scheme:

Premise 1.	X1, an explanation of Y, meets criteria C1 – Cn with scores S1 – Sn
Premise 2.	Decision rule R applies.
Conclusion.	Of possible explanations X1 – Xn of Y, X1 is the best explanation of Y.

Notice that I've modified the first premise slightly in this reconstruction, to include a reference to Y, the observation to be explained. Otherwise, the premises would seem under constrained. The conclusion is about explanations of Y, not explanations of anything.

One thing about these schemes that seems somewhat inadequate as an account of abduction in the sense of a method for finding the best explanation of an observation, inherited from Wagemans' original formulation, is that alternative explanations are not explicitly compared with X1. The other explanations, X2 to Xn, are referenced, but the comparison of their scores with the scores of X1 are buried in the decision rule, R, and not transparent. Here I think there may be some room for improvement.

REFERENCES

Besnard, P., & Hunter, A. (2008). *Elements of Argumentation*. MIT Press.

Poole, D. (1988). A logical framework for default reasoning. *Artificial Intelligence*, *36*(1).

Prakken, H. (2010). An abstract framework for argumentation with structured arguments. *Argument & Computation*, *1*, 93–124.

Walton, D. (1996). *Argumentation Schemes for Presumptive Reasoning*. Mahwah, NJ: Lawrence Erlbaum Associates.

Walton, D. (2006). *Fundamentals of Critical Argumentation*. Cambridge: Cambridge University Press,
Walton, D., Reed, C., & Macagno, F. (2008). *Argumentation Schemes*. Cambridge: Cambridge University Press.