Inductive Generalizations are Abductions

Position paper for the Workshop on Abductive and Inductive Reasoning
(Adapted from Josephson & Josephson, 1994. Used with permission.)

John R. Josephson
The Ohio State University
Laboratory for AI Research
Computer and Information Science Department
Columbus, Ohio, USA

What is abduction?

What I mean by "Abduction" or "Inference to the Best Explanation" is a form of inference that follows a pattern like this:

D is a collection of data (facts, observations, givens),

H explains D (would, if true, explain D),

No other hypothesis explains D as well as H does.

Therefore, H is probably correct.

The strength of an abductive conclusion will in general depend on several factors, including:

* how good H is by itself, independently of considering the alternatives,
* how decisively H surpasses the alternatives, and
* how thorough the search was for alternative explanations.

When I say that the strength of abductive conclusion "will in general" depend on these factors I mean that it should depend on these factors, and that insofar as we are intelligent creatures, our conclusions will actually depend on these factors. Thus these factors are descriptive as well as normative considerations.

It is useful to distinguish abduction as a pattern of justification (given in the schema above), from abduction as a reasoning process. An abductive reasoning process aims at a satisfactory explanation, one that can be confidently accepted. An explanation that can be confidently accepted is one that has a strong abductive justification. Thus, abductive reasoning aims at setting up abductive justifications.

The abductive justifications set up by abductive reasoning might be explicit, as when a diagnosis can be justified, or they might arise implicitly as a result of the functioning of an "abductively effective" mechanism, such as, perhaps, the human visual system or human language understanding.

This paper is about justification, not reasoning processes. For a treatment of abductive reasoning, especially the computational challenges in forming well-justified composite explanatory hypotheses, see Josephson & Josephson (1994).

Inductive Generalizations are Abductions

An inductive generalization is an inference that goes from the characteristics of some observed sample of individuals to a

conclusion about the distribution of those characteristics in some larger population.[1]

I will argue that it is possible to treat every good (i.e., reasonable, valid, strong) inductive generalization as an instance of abduction, and that analyzing inductive generalizations as abductions clarifies evaluating the strengths of these inferences.

As Harman (1965) pointed out, it is useful to describe inductive generalizations as abductions because it helps to make clear when the inferences are warranted. Consider the following inference:

All observed *A* 's are *B* 's

_____

Therefore All *A* 's are *B* 's

This inference is warranted, Harman writes, ". . . whenever the hypothesis that all *A* 's are *B* 's is (in the light of all the evidence) a better, simpler, more plausible (and so forth) hypothesis than is the hypothesis, say, that someone is biasing the observed sample in order to make us think that all A's are B's. On the other hand, as soon as the total evidence makes some other competing hypothesis plausible, one may not infer from the past correlation in the observed sample to a complete correlation in the total population."

If this is indeed an abductive inference, then "All *A* 's are *B* 's" should explain "All observed *A* 's are *B* 's." But, "All *A* 's are *B* 's" does not seem to explain why "This *A* is a *B* ," or why *A* and *B* are regularly associated (pointed out by Ennis, 1968). Furthermore, in general we usually think that explanations give causes[2], but it is hard to see how a general fact could explain its instances, because it does not seem in any way to cause them.

The story becomes clearer if we are careful about what precisely is explained and what is doing the explaining. What the general statement in the conclusion explains are characteristics of the set of observations, not the facts observed. For example, suppose I choose a ball at random (arbitrarily) from a large hat containing colored balls. The ball I choose is red. Does the fact that all of the balls in the hat are red explain why this particular ball is red? No, but it does explain why, when I chose a ball at random, it turned out to be a red one (because they all are). "All *A* 's are *B* 's" cannot explain why "This *A* is a *B* " because it does not say anything at all about how its being an *A* is connected with its being a *B* . The information that "they all are" does not tell us anything about why this one is, except that it suggests that if we want to know why this one is, we would do well to figure out why they all are.

A generalization helps to explain some characteristics of the set of observations of the instances, but it does not explain the instances themselves. That the cloudless, daytime sky is blue helps explain why, when I look up, I see the sky to be blue (but it doesn't explain why the sky is blue). Seen this way, an inductive generalization does have the form of an inference whose conclusion explains its premises.

In particular, "*A* 's are mostly *B* 's" together with "This sample of *A* 's was obtained without regard to whether or not they were *B* 's" explain why the *A*'s that were sampled were mostly *B* 's.

> Why were 61% of the chosen balls yellow?
>
> Because the balls were chosen more or less randomly from a population that was two thirds yellow (the difference from 2/3 in the sample being due to chance).

Alternative explanation for the same observation:

> Because the balls were chosen by a selector with a bias for large balls from a population that was only one third yellow but where yellow balls tend to be larger than non yellow ones.

<u>The frequencies in the larger population, together with the frequency-relevant characteristics of the method for drawing a sample, explain the frequencies in the observed sample.</u>

What is explained? In this example, just the frequency of characteristics in the sample is explained, not why these particular balls are yellow or why the experiment was conducted on Tuesday. In general, the explanation explains why the sample frequency was the way it was, rather than having some markedly different value. If there is a deviation in the sample from what you would expect, given the population and the sampling method, then you have to throw some Chance into the explanation (which is more or less plausible depending on how much Chance you have to suppose).

How are frequencies explained? Frequencies are explained by giving a causal story that explains how the frequencies came to be the way they were. This causal story typically mentions both the population frequency and the method of drawing the sample. Unbiased sampling processes tend to produce representative outcomes; biased sampling processes tend to produce unrepresentative outcomes. This "tending to produce" is causal and supports explanation and prediction. A peculiarity is that characterizing a sample as "representative" is characterizing the effect (sample frequency) by reference to part of its cause (population frequency). Straight inductive generalization (carrying the sample frequencies unchanged to the generalization) is equivalent to concluding that a sample is representative, which is a conclusion about its cause. Inductive generalization depends partly on evidence or presumption that the sampling process is (close enough to) unbiased. The unbiased sampling process is part of the explanation of the sample frequency, and any independent evidence for or against unbiased sampling bears on its plausibility as part of the explanation.

If we do not think of inductive generalization as abduction, we are at a loss to explain why such an inference is made stronger or more warranted, if in collecting data we make a systematic search for counter-instances and cannot find any, than it would be if we just take the observations passively. Why is the generalization made stronger by making an effort to examine a wide variety of types of *A* 's? It is made stronger because the failure of the active search for counter-instances tends to rule out various hypotheses about ways in which the sample might be biased, that is, it strengthens the abductive conclusion by ruling out alternative explanations for the observed frequency.

References

Josephson, J. R., & Josephson, S. G. (Eds.). (1994). *Abductive Inference: Computation, Philosophy, Technology*. New York: Cambridge University Press.

Harman, G. (1965). The Inference to the Best Explanation. *Philosophical Review*, 74, 88-95.

Ennis, R. (1968). Enumerative Induction and Best Explanation. *The Journal of Philosophy*, LXV(18), 523-529.

---

[1]Notice that so-called "concept learning from examples" (supervised or unsupervised) is a family of kinds of inductive generalization. The learned concept generalizes the frequencies of occurrence and co-occurrence of certain characteristics in a sample with the intention to apply them to a larger general population, which includes unobserved as well as observed instances.

Another inference form that has often been called "induction" is given by the following:

All observed *A*'s are *B*'s.

Therefore, the next *A* will be a *B*.

Let us call this inference form an "inductive projection." Such an inference can be analyzed as an inductive generalization followed by a prediction, as follows:

Observations ----> All *A*'s are *B*'s ----> The next *A* will be a *B*.

[2] In Josephson and Josephson (1994, p. 16 ff.) I argue that some good explanations are not proofs and that some proofs are not explanations. I suggest that representing explanations as proofs tends to confuse causation with implication; implications occur in the mind, causation occurs in the world. Instead, a better view is that explanations assign causal responsibility, but that causality must be understood more broadly than the usual senses of mechanical, efficient, or event-event causation.