Acta Analytica (2024) 39:163–182
https://doi.org/10.1007/s12136-023-00555-2

Deduction, Abduction, and Creativity

Tomáš Hanzal[1]

Received: 17 October 2022 / Accepted: 3 April 2023 / Published online: 19 April 2023

Abstract
In a discussion of Sherlock Holmes' "science of deduction" and the related "method of exclusion," I show that Holmes' claim that his inferences are deductive makes sense, if we consider his theoretical presuppositions. So, it is more accurate to say that he tries to reduce abduction to deduction than that he confuses them. His theoretical framework, albeit inadequate as a theory of empirical reasoning, can be seen as a basic model of classical (symbolic) AI. The main problems of this approach are surveyed, and abduction is brought into play as both a better characterization of Holmes' inferences and a better guide for building AI systems. This is a good background for raising the question of creativity because, according to Peirce, it relates to abduction in a substantial way. Is Sherlock Holmes creative? Can machines be creative? An affirmative answer to the second question might be given by a "logic of discovery" but the problem remains that inventing such a logic already presupposes creativity.

Keywords Deduction · Abduction · Creativity · Sherlock Holmes · Artificial intelligence · Logic of discovery

1 Introduction

The aim of this paper is to examine some connections between *deduction* as conceived by Sherlock Holmes, *abduction* as conceived by C. S. Peirce, and *creativity* insofar it can be seen in some systems of *symbolic* artificial intelligence (AI). What connects those in the first place is the topic of *searching for explanation of occurrences*, ranging from "everyday reasoning" to discovery of scientific laws.

The starting point are the stories of Sherlock Holmes, which have the reasoning involved in searching for explanations for one of their main themes. The occurrences that Holmes strives to explain are primarily crimes, but he also demonstrates his

✉ Tomáš Hanzal
tomas.hanzal@ff.cuni.cz

[1] Institute of Philosophy and Religious Studies, Faculty of Arts, Charles University, Prague, Czech Republic

reasoning in everyday matters, such as explaining how it happened that Watson's shoes are dirty. The reason for choosing Arthur Conan Doyle's stories as the starting point, however, is not just that Holmes provides examples. He also has a particular theory about the nature of his inferences. Specifically, he claims that they are *deductions*. However, this does not seem to be right, as has been shown by scholars, usually with reference to Peirce's theory of *abduction*. This is reviewed in Sect. 2.

I agree that Holmes' typical reasonings are best explained as abductions, but I disagree that Holmes would be simply mistaken to talk about "deduction." In Sect. 3, I show that according to his theory, under ideal conditions, abduction would be *reducible to deduction*. However, in real-life reasoning, this reduction is not practicable because of the unwarranted *closed-world assumption* that Holmes' theory implicitly makes (as shown in Sect. 4). The analysis of Holmes' theory of reasoning also shows that it can also be seen as a simple model of an AI system (a "reasoning machine" or an "ideal reasoner").

The idealized reduction of abduction to deduction is not practicable even for AI but, as it turns out, there are AI systems that explicitly make use of *abduction* (as illustrated in Sect. 5). This opens the possibility of an "ideal reasoner" Holmes would like to see, but without unrealistic restrictions. But there appears another problem. Unlike Holmes (and other people), a machine of this type would need much more *explicit* knowledge — "knowledge" that we would normally classify as implicit or "tacit" (if we would classify it as knowledge at all). In Sect. 6, I claim that at least *some types* of abduction depend on such "tacit" knowledge and also that such abductions can be the source of (explicit) knowledge that further abductions depend on (e.g., scientific laws).

Such abductions can be called *creative*. In Sect. 7, several typologies of abduction are briefly surveyed with emphasis on how they view *creative* abductions and whether Holmes' abductions can be viewed as creative. The answer is ambiguous. The individual steps of his abductive inferences may be very mundane, but when taken together, "chains" of them could lead to unconventional and novel conclusions, hence comes the idea of reducing creativity to combinatorics, which is related to the idea of a *logic of discovery* (Sect. 8). Such logic would allow us to do things that are considered creative algorithmically. This means that an "ideal reasoner" could be creative. There is a catch, however, because implementing a logic of discovery usually requires some *heuristics*, and these heuristics are arguably products of creativity, too.

2 "The Book of Life," Deduction and Abduction

Sherlock Holmes is famous, among other things, for his so-called science of deduction. The idea behind it is, as he puts it in his article *The Book of Life* quoted by Watson, that "all life is a great chain, the nature of which is known whenever we are shown a single link of it."[1] That means that everything in life, or nature, is connected

[1] *A Study In Scarlet*, p. 13. The stories of Sherlock Holmes are quoted from Doyle (2013).

in a predictable, law-like fashion and if we know the laws, i.e., if we can "read in the book of life," we can *deduce anything*. We can, as his example goes, tell anyone's history and profession from a glance, if we only observe the right signs such as fingernails, coat sleeves, and boots. Similarly, from the inspection of a crime scene, we should be able to tell exactly what events lead to the crime. This seems a little incredible, and indeed, the perfect exercise of the "science of deduction" is just a limiting case because, as Holmes tells Watson (and us), it would require the "possession of all knowledge."[2] That is what keeps us mortals from being able to deduce *everything* that is in any way connected with what we observe. Nonetheless, according to Holmes, "detection is, or ought to be, an exact science."[3] The fact that no mortal has "all knowledge" does not imply that it would not be theoretically possible to deduce from observed facts in the way Holmes imagines (and imagines himself doing in cases where his knowledge is complete enough).

However, a distinction that is reflected in Holmes' theory of reasoning shows us what could be the decisive disanalogy between exact science and Holmes's detection (or reading the book of life). It is the distinction between, as he calls it, "synthetic" (forward) and "analytic" (backward) reasoning. The former proceeds from a "train of events" to their *result*, and the latter, which seems to be necessary for detection, but, according to Holmes, is rare in everyday life, proceeds from the result to the preceding events.[4] The problem now is that to the former kind belong not just our everyday reasonings but, apparently, also those that could be called *deductions* of exact sciences. The role of deduction in science mainly consists in drawing consequences of laws or theorems. If there is some sense of "deduction" relevant to natural science and at the same time related to Holmes' usage of this word, it is *prediction*. Science gives us laws that allow us to predict something from some initial conditions. Holmes, who is not only a practicing detective and a theoretician of reasoning but also a detection scientist, offers us some "laws" analogous to those of natural science. Holmes confesses that he has written "several monographs," such as "Upon the Distinction between the Ashes of the Various Tobaccoes" and one "upon the influence of a trade upon the form of the hand, with lithotypes of the hands of slaters, sailors, corkcutters, compositors, weavers, and diamond-polishers."[5]

With the knowledge we could presumably find in Holmes' books, it would be possible to deduce, for example, that smoking such-and-such kind of tobacco (under some idealized conditions) will produce such-and-such ashes. But the point of his writings is being able to use the knowledge not for prediction, but in the opposite direction. This is possible, of course. We can *infer* what kind of tobacco was smoked from what kind of ashes we have found, what is somebody's occupation from the deformation of their fingers, or what animal has crossed our path from the traces it has left. But these inferences will not, standardly, count as *deductions* because they are not *necessary* (not even under ideal conditions). Holmes knows that certain

[2] *The Five Orange Pips*, p. 229.
[3] *The Sign of the Four*, p. 88.
[4] *A Study In Scarlet*, p. 81.
[5] *The Sign of the Four*, p. 89. An overview of Holmes' writings can be found in Klinefelter (1989).

tobacco produces such-and-such ashes, but that does not imply the validity of the converse implication. Assuming the converse implication would actually be a logical fallacy. Similarly, in the case of hand deformation caused by one's occupation, the implication in the "laws" or rules goes in the direction from the occupation to the deformation and not in the opposite direction. This is evident in Holmes' own practice.

In *The Solitary Cyclist*, a client comes to Holmes, who confesses that he nearly supposed mistakenly that she was a typist, but, in the end, he correctly inferred that she was a musician. The thing is that, according to his knowledge, both professions are causes of "spatulate finger-ends" — but typewriting does not cause "a spirituality about the face."[6] This is also an example that shows how everyday reasoning concerning empirical phenomena significantly differs from formal deduction: It is *nonmonotonic*, meaning that adding another premise can change the validity of an inference. The material inference.

x has spatulate finger-ends $\rightarrow x$ is a typist
seems valid (not necessarily, of course), but
x has spatulate finger-ends $\wedge$ x has a spirituality about the face $\rightarrow x$ is a typist

seems invalid, according to Holmes' somewhat peculiar "knowledge." Moreover, Holmes did not actually *deduce* anything here, and the fact that he was right was in part by chance. According to his knowledge, typing does not produce "spirituality about the face," but that obviously does not mean that he could deduce that someone with this spirituality (whatever it is) is not a typist because it could be caused by something not related to the occupation.

It is no wonder that Holmes' claims about *deduction* have been disputed for decades. It has been repeatedly pointed out that the most common type of his inferences is not deduction (nor induction), but *abduction*, the third type of inference or argument made famous by Charles Sanders Peirce. The connection between Sherlock Holmes and abductive reasoning was mentioned already by Fann (1970, pp. 57–58). Abduction and reading most of Holmes' arguments as abductive are the common themes of several articles in a volume edited by Umberto Eco and Thomas A. Sebeok entitled *The Sign of Three: Dupin, Holmes, Peirce* (1983).

Peirce defines abduction in several ways, but they basically match Holmes' conception of "reasoning backwards"[7] or "analytic" reasoning,[8] i.e., reasoning from the *result* to something that would explain it, or from effect to cause (Peirce, 1878, p. 477). Or, as Peirce puts it elsewhere:

The surprising fact, C, is observed;
But if A were true, C would be a matter of course,
Hence, there is reason to suspect that A is true. (Peirce, 1974, 5.189)

[6] *The Solitary Cyclist*, p. 652.

[7] Cf. Peirce's term "*retroduction*" — see Peirce 1974, 1.65; Sebeok and Umiker-Sebeok (1983), p. 39.

[8] Note that for Peirce "*analytic*" inference means, on the contrary, *deduction*. Induction and hypothesis are kinds of *synthetic* inference (Peirce 1878, p. 472). Holmes' terminology may be inspired by Greek geometry (see Magnani 2001, p. 2).

But the difference between deduction and abduction is not just the "direction." Deduction is necessary,[9] whereas abduction is probable at best. As Peirce says, it is a "weak kind of argument"; "we only surmise that it may be so" (Peirce, 1878, p. 473). These *might be* the hands of a typist but also of a musician. Therefore, if we say that Holmes is wrong about the nature of his reasonings, that they are not deductive but, in fact, abductive, the critique is not just on the level of terminology. It is surely possible to dismiss this issue by taking Holmes' usage of "deduction" to mean "inference" in general, and, as Massimo Pigliucci (2012), attribute this mistake to Holmes' ignorance of philosophy that is mentioned by Watson.[10] But that does not seem entirely right. Holmes really wants us to see his method as thoroughly scientific and his inferences as necessary. In many examples of his reasonings, this seems doubtful. But can his insistence on deduction be made sense of? That is what I will look at before returning to abduction.

3 Deductive Method of Elimination

We have already seen that there is one somewhat stringent condition for the "ideal reasoner" — the "possession of *all knowledge*."[11] I take this to be everything that science could tell us about the laws governing the connections between things and events in nature — including, for example, the connection between doing a certain job and deformations of the finger-ends. Holmes is obviously an expert in this area, but his knowledge could surely be even better. But even if he had the knowledge that would enable him to distinguish a musician's fingertips from a typist's at first sight, it would not enable him to make a *deduction*. The "knowledge" would still have this form:

x is a typist $\rightarrow$ x has so-and-so deformed finger-ends

Its direction is from a "cause" to its effect or "result." But to enable a *deduction* (from a result to its cause), the law would have to be the *converse* of this implication. In other words, it would have to say that such-and-such occurrence *can have only one determinate cause* and no other. I do not claim that this is not theoretically possible. There are scientific laws in the form of biconditionals, which support deduction in both directions. Holmes' ideal knowledge could be of this form:

x is a typist iff x has so-and-so[12] deformed finger-ends

The first conditional would correspond to reasoning "forward" (meaning from cause to effect), the second to reasoning "backward." Let's say that having the knowledge corresponding to the second conditional is possible (it might be, in principle, attained statistically). The problem is that Holmes does not actually have

[9] Leaving aside *statistical* deduction, which was also something of interest to Peirce. See Levi (2006).
[10] *A Study In Scarlet*, p. 12.
[11] *The Five Orange Pips*, p. 229.
[12] This "so-and-so" would have to be more specific than in the conditional above.

knowledge of this form, as the example with the supposed typist shows, but performs what he calls "backward" reasoning nonetheless.

There is, however, another way to turn Holmes' inferences into deductions. It involves his *method of exclusion* (or *elimination*), which is formulated multiple times in the stories over the period of many years. One of the famous formulations goes like this:

> It is an old maxim of mine that when you have excluded the impossible, whatever remains, however improbable, must be the truth.[13]

It is often the case that, according to Holmes' knowledge, some fact has multiple possible explanations, multiple possible causes. We have already seen that we should blame rather the form of the knowledge than its want for this fact. So let us say that Holmes has a "surprising fact" C to explain. In his "knowledge base," he has these relevant causal "laws":

$A \rightarrow C$
$B \rightarrow C$

In general, any combination of the truth values of A and B is possible. But if we assume this is *all* (relevant) knowledge and that there must be a *sufficient reason* for C, then we can *deduce:*

$C \rightarrow A \vee B$[14]

Under these suppositions, the method of exclusion can be applied. Now, if we can eliminate either A or B (or all but one if there are more than two possibilities) using deduction (e.g., by incompatibility with other evidence), then we prove the reality of the other (or the one remaining) possibility *deductively*.

4 Problems of a Closed World

This is how I think Holmes' talk about deduction *can* be given sense. The problem, of course, is that there is another supposition that is needed for the method of exclusion to be practicable: that the world is *closed* in the sense that we can *enumerate everything*, i.e., for example, that we can have a finite set of all possible explanations (see Bonfantini and Proni (1983), pp. 127–128; Hintikka and Hintikka (1983), p. 160f.; Smajić, 2010, pp. 126–128). The supposition of "all knowledge" is equivalent to the *closed-world assumption* with respect to natural laws (in a wide sense) — *law-like statements that we do not know to be true are false*. Making even this assumption is problematic, but Holmes must go even further. To make the method

[13] *The Beryl Coronet*, p. 324. Other occurrences: *A Study In Scarlet*, p. 82; *The Sign of the Four*, pp. 89, 111; *The Bruce-Partington Plans*, p. 968; *The Blanched Soldier*, p. 1054.

[14] A and B may or may not be compatible. It seems that Holmes also presupposes that one of them holds exclusively (for example, if the perpetrator escaped through the window, we presume he did not escape through the door as well). But this is not relevant to the method of exclusion itself.

of exclusion work, he must at some point make this assumption even with respect to evidence, which is, I would say, scandalous.

Say there were several suspects in a murder case. Holmes managed to show that it is impossible that any of them, except one, committed the murder. But there can also be someone else, who Holmes does not know of. Holmes claims repeatedly that he came to the right conclusion using the method of exclusion, which might be right, but also a piece of luck at the same time. It shows us only that he usually managed to think of the right solution among some others, not that he really thought of *all* possibilities. So, the main problem of Holmes' theory of reasoning is not that he cannot distinguish deduction from abduction, but that he sees knowledge as static and the world, or at least, the sets of possible explanations as finite or closed. If this was right, he would also be right about deduction. In other words, abduction is in Holmes' dream-world reducible to deduction.

This goes in line with the metaphor of a *reasoning machine* that is used several times by Watson.[15] He calls Holmes a machine, especially because of his occasional or apparent lack of emotion, but the idea that Holmes' "ideal reasoner" could actually be a machine, a computer, also plays its role. (Of course, it would have to be not just a reasoning machine but also an *observing* machine, which complicates matters further, but this is not the topic of this paper.)

So, Holmes' ideal reasoner is also an ideal of artificial intelligence. Given the closed-world assumption, we can imagine Holmesian reasoning as computing. Unfortunately, this is not practicable. We have to give up deduction in Holmes's strong sense.

The first concession is marked already in Holmes' later formulation of the method of exclusion:

> […] when you have eliminated all which is impossible, then whatever remains, however improbable, must be the truth. It may well be that several explanations remain, in which case one tries test after test until one or other of them has a convincing amount of support.[16]

So, Holmes noticed that, in reality, he cannot always eliminate all "possible" explanations except one as impossible (i.e., "deductively"). The words "convincing amount of support" suggest that he also has to work with plausibility or *probability* (as suggested also by his remarks elsewhere).[17] But this does not change anything about the problematic closed-world assumption with respect to candidate explanations (and, therefore, to evidence).

The closed-world assumption with respect to "knowledge" is something we can make for the purpose of some knowledge-based systems of artificial intelligence working in some clearly defined domain, but detection encompasses too much for this to be practicable. A good example of a narrower domain is medicine and the

[15] *The Sign of the Four*, p. 94; *A Scandal in Bohemia*, p. 163; *The Crooked Man*, p. 424; *The Six Napoleons*, p. 720.

[16] *The Blanched Soldier*, p. 1054.

[17] *The Hound of the Baskervilles*, p. 516.

reasoning involved in *differential diagnosis*, as shown in the Holmes-inspired TV series *House*.[18] House's world, like Holmes', seems to be *closed* with respect to "knowledge." He and his colleagues do not discover new diseases or unknown symptoms of old diseases. Differential diagnosis is, in fact, the same thing as Holmes' method of exclusion. We see some "symptoms," and we have knowledge that links these symptoms to (all) their possible causes. A process of elimination has to take place. It involves looking for more evidence (not just other symptoms and results of medical tests, but also evidence, for example, of using drugs in the patient's apartment, if they are among possible causes), and of course also inferences including deductive ones. Often, the chosen diagnosis must be verified or falsified according to the patient's reaction to the treatment that is appropriate for it.

So, even with the presupposition that we have *all knowledge*, not just knowledge of all (relevant) laws but also of all possible explanations, of which one must be true, the elimination of "the impossible" is only partially a deductive matter.

5 Automatic Abduction

Now, the fact that neither Sherlock Holmes' detective reasoning nor differential diagnostics can actually be characterized as deductions does not mean that a computer cannot do it in principle. In fact, there are knowledge-based systems of artificial intelligence whose purpose is exactly this, finding possible causes/explanations of some occurrences, i.e., making abductive inferences.[19] In the field of AI, this is usually called *backward chaining*, which corresponds to Holmes' "reasoning backwards" and Peirce's "retroduction." We can use as an example the "scheme for abductive inference" developed in the TACITUS project (Hobbs et al., 1993). I will not go into the technical details, but there are a few interesting points to mention.

The authors say explicitly that the kind of inference the system does is *abduction* (already in the title of the paper: *Interpretation as Abduction*), even though they characterize it as "*inference to the best explanation*," which is pretty standard in many contexts but not something that I would endorse here.[20] The purpose of the system is to *interpret* sentences, which, according to the authors, "can be viewed as the process of providing the best explanation of why the sentences would be true." This might look like a task that is a bit different from the kind of reasoning we know from Holmes but, in principle, it isn't. The (possible) verity of the sentence is an occurrence that needs an explanation, just like a trace or a symptom. The idea that interpretation involves abduction, or making *hypotheses*, is also acceptable. Grice's "conversational implicatures," for example, which are mentioned by the authors, can be seen as such abductive inferences (Hobbs et al., 1993; cf. Grice, 1975).

[18] *House M. D.*, 2004—2012.

[19] See, e.g., Magnani (2001) for an example from the mentioned medical domain.

[20] I use Peirce's characterizations of abduction. Also, the notion of the *best* explanation is problematic (see Sect. 7).

However, our problem now is not specifically interpretation but abduction insofar as it can be carried out algorithmically. *Expert systems*, such as the one mentioned, are basically made up of two parts, the *inference engine* and the *knowledge base*.[21] The scheme for abductive inference, which drives the inference engine in this case, is far from being trivial and utilizes weights and costs (which is something people, including Holmes, usually do not do — at least not explicitly). But what we are interested in now is the *knowledge base*. The general point about algorithmic abduction is that it needs *a lot* of knowledge (ideally, as we know from Holmes, "all knowledge") *in explicit form*.[22] And it does not require just the more-or-less scientific knowledge Holmes has in mind, but also things that we would not normally call "knowledge" at all because they are rather parts of our (linguistic) *know-how* or our *tacit knowledge*.[23] Let us take an example from the TACITUS paper:

> The Boston office called. (Hobbs et al., 1993, p. 71)

The system must be able to "abduce" that there is *some person* who called (because only a person may call, not literally an office) that has some appropriate relation to the office *in* Boston. Say, we further have in our knowledge base that there is a particular person who works for the office. Then, we must also *know* that "works-for" is a relation. And so, if this person called, the original interpreted sentence is true.

In dealing with various pragmatic linguistic phenomena, still more axioms must be added (in a seemingly ad hoc fashion). For example:

$$\text{doctor(d)} \rightarrow \text{person(d)}$$
$$\text{person(d)} \wedge \text{male(d)} \rightarrow \text{he(d)}$$

Backward chaining these axioms allows the system to determine that "he" in one sentence may anaphorically refer to "doctor" in the previous (pp. 91–92).[24]

This example may be tied specifically to the problem of interpretation of sentences, but the point is general: Holmes' "ideal reasoner" conceived as a "machine" would need to have much more than "*all* knowledge" (as conceived by Holmes) in its knowledge base.

Let us consider an example from Sherlock Holmes. When Holmes and Watson met for the first time, Holmes immediately "perceived" that Watson had been in Afghanistan. Watson was, as in many other similar cases during many subsequent years, astonished by Holmes' quick conclusion, but deemed the reasoning involved "simple enough" after Holmes explained it to him later:

> I knew you came from Afghanistan. From long habit the train of thoughts ran so swiftly through my mind, that I arrived at the conclusion without being con-

[21] Cf. Holmes' three qualities of ideal detective, which include "power of deduction" and knowledge (*The Sign of the Four*, pp. 88–89).

[22] See Dreyfus (1992).

[23] See Polanyi (2009). Regarding "tacit information" with respect to the interpretation of Sherlock Holmes, see Hintikka and Hintikka (1983).

[24] I use simplified notation in the axioms, leaving out the universal quantifiers.

scious of intermediate steps. There were such steps, however. The train of reasoning ran,

(1) 'Here is a gentleman of a medical type, but with the air of a military man. Clearly an army doctor, then.
(2) He has just come from the tropics, for his face is dark,
(3) and that is not the natural tint of his skin, for his wrists are fair.
(4) He has undergone hardship and sickness, as his haggard face says clearly.
(5) His left arm has been injured. He holds it in a stiff and unnatural manner.
(6) Where in the tropics could an English army doctor have seen much hardship and got his arm wounded? Clearly in Afghanistan.'

The whole train of thought did not occupy a second.[25]

I added the numbering to the individual *abductive* inferences (cf. Genot, 2020), but Holmes does so himself in a similar case elsewhere.[26] This explanation (or explication) might make it "simple enough" for Watson, but not yet for the algorithmic reasoner. A lot of not exactly specialist "knowledge" would be required in *explicit* form to support the abductive inferences. To mention just a few simple examples (and leave more to the reader):

$$\text{has-been-in-the-tropics}(x) \rightarrow \text{tanned}(x)$$
or
$$\text{has-injured-arm}(x) \rightarrow \text{holds-arm-in-a-stiff-and-unnatural-manner}(x)$$

And even these "axioms" (especially the last one) are questionable. Having one's arm injured does not *necessarily* lead to holding it in a stiff manner (just as holding one's arm in a stiff manner can have other causes). This can be fixed using an interesting device from the TACITUS project — the "et cetera" propositions. These allow us to formulate axioms that are not completely explicit (but they explicitly acknowledge this). In this example, it would be[27]:

$$\text{has-injured-arm}(x) \wedge \text{etc}_1(x) \rightarrow \text{holds-arm-in-a-stiff-and-unnatural-manner}(x)$$

which can be read as: Who has an injured arm holds it in a stiff and unnatural way, unless [not $\text{etc}_1(x)$]. This provides a way to acknowledge the *nonmonotonicity* of common-sense reasoning (Hobbs et al., 1993, p. 85) while allowing the use of incomplete or provisional knowledge in abduction. So, unlike Holmes' "science of deduction," algorithmic abduction does not need the closed-world assumption with respect to knowledge. It needs neither *all* knowledge, nor axioms encompassing all possibilities. But all knowledge that it can use in abductive inference must be explicit, and "knowledge" in this sense includes each and every triviality.

[25] *A Study In Scarlet*, p. 14.
[26] *The Dancing Men*, p. 634.
[27] Cf. Hobbs et al. (1993), pp. 84–87. There, it is used for axioms concerning classification. For example: $\text{mammal}(x) \wedge \text{etc}_2(x) \rightarrow \text{elephant}(x)$. Thanks to this, being a mammal can serve as a (somewhat weak, of course) evidence for being an elephant.

6 Abduction and Explicit Knowledge

So far, we have seen the following.

(1) Sherlock Holmes often characterizes his inferences as "deductions," which raises justified doubt, but
(2) His model of his inferences as deductive can be made sense of — with the help of the *method of exclusion* and the *closed-world assumption* with respect to both knowledge and evidence.
(3) However, this is something we are not justified in assuming, especially in Holmes' detective domain.
(4) That, on the other hand, does not mean that we have to give up Holmes' vision of the "ideal reasoner" conceived as a reasoning machine, basically a computer algorithm, a backward chaining, or *abductive*, inference engine.
(5) Such an "ideal reasoner," however, differs from people in that respect that its *knowledge base* must *explicitly* contain all knowledge that can be used for inferences — so, in a sense, much more than Holmes' "all knowledge."

Now, this is not to say that the "ideal reasoner" cannot, in principle, have *all the knowledge* required for reasoning (in a specific domain) much like a human expert. Nor that computers cannot, in a sense, have something that could be called "implicit" knowledge.[28] But much that would have to be explicitly encoded knowledge of the reasoning machine does not need to be explicit either for (e.g.) Holmes to be able to use it in his reasoning or for the conclusion of the reasoning (i.e., the explanation) to be comprehensible and plausible. This "knowledge," then, can, of course, be *made explicit*, but it does not need to be. This leads us to two related questions: (1) Can abduction generally be conceived as based only on (already) explicit "knowledge"? and (2) Where does this explicit knowledge come from?

It looks like Holmes' ideal is just to do reasoning and avoid (creative) *thinking*. He has the knowledge, and he has the algorithm (be it deductive or abductive). With this, he can theoretically solve any problem. There are differences in complexity, but those can be quantified: "It is quite a *three pipe* problem,"[29] Holmes says to Watson to indicate that he needs some time for his reasoning. But even if it were so, there must be the "knowledge base," which he must have somehow acquired. For Holmes, this probably means mainly reading scientific books and encyclopedias[30] but partly, as already mentioned, also his own research, which should make it clear even to him that "all knowledge" is by no means "static," unchangeable, or possibly once forever

[28] In the *connectionist* paradigm (as opposed to the "classical" or symbolic that is presupposed here), the structure and weights of a neural network can be seen as containing implicit knowledge. The problem, in that case, is how to make it explicit and use it in an explanation. *Interpretability* is a notorious problem of neural networks, and an important subject of research. See, for example, Fan et al. (2021).
[29] *The Red-Headed League*, p. 187, emphasis mine.
[30] See *The Five Orange Pips*, pp. 229–230.

complete. In other words, there is always room for *new ideas*, and without ideas that were at some point new, we wouldn't have much knowledge.

Peirce says that neither induction nor deduction "can originate any idea whatever," and:

> All the ideas of science come to it by the way of Abduction. Abduction consists in studying facts and devising a theory to explain them. Its only justification is that if we are ever to understand things at all, it must be in that way. (Peirce, 1974, 5.145)

So, abduction is the type of reasoning involved not only in detective reasoning (which also matches the description given by Peirce: "studying facts and devising a theory to explain them") but also in devising scientific theories and therefore, arguably, the scientific knowledge Holmes used in his abductions.

A simple example of the "scientific" abduction might be the idea Holmes must have had at some point: Tobacco ashes do not always look the same — that would be explained if different kinds of tobacco produced *regularly* different kinds of ashes. Peirce's favorite example of abduction in science is Kepler's discovery that the orbit of Mars has the shape of an ellipse with the sun in one of its focal points. Peirce even calls it "the greatest piece of Retroductive reasoning ever performed" (Peirce, 1974, 1.74). The thing is that the hypothesis, that the orbit is elliptical rather than circular, although it may seem "simple enough" now, was by no means obvious at the time. It was not a simple extrapolation from the data, and it was certainly not just choosing the ellipse from some predefined short-list of possible shapes. Kepler had to go through many steps and hypotheses before reaching the final one. But every modification of the theory was logical in the sense that it led to better explanation of the observations.[31]

This aspect of abduction, that it can be a source of new ideas, is obscured when it is presented as an *inversion* of deductive *syllogism*, which is what Peirce did in his earlier texts (where he did not use the term "abduction" yet and was talking about "*hypothesis*"). While deduction is, according to Peirce of these texts, inference from *rule* and *case* to *result*, and induction from *case* and *result* to *rule*, hypothesis is an inference from *rule* and *result* to *case*. Thus, his famous example:

> *Rule*.—All the beans from this bag are white.
> *Result*.—These beans are white.
> *Case*.—These beans are from this bag. (1878, p. 472)

The problem is that this makes it look like we *already have* the *rule* at the outset. Well, according to the Holmesian/AI model, we do, *not as a premise*, however, but as one of many items in the knowledge base (potentially along with some alternatives that could play the role of the *rule* in the syllogism as well). Nancy Harrowitz puts this problem in terms of the "chronology of information-obtaining" — only the *result* (i.e., the observed fact) is first and

[31] See Peirce 1974, 1.71–73; Peirce 1974, 2.96; Hanson 1958, pp. 72–85.

> The process of abduction takes place between the result and the rule, and concludes with the positing of a hopefully satisfactory hypothesis. (Harrowitz, 1983, pp. 182–183)[32]

Umberto Eco makes a similar point when he says that

> the real problem is not whether to find first the Case or the Rule, but rather how to figure out both the Rule and the Case at the same time, since they are inversely related, tied together by a sort of chiasmus. (Eco, 1983, p. 203)

So, the *rule* (and the *case* with it) must be "found" in the process of abduction, which is, of course, what the "inference engine" does, but sometimes it must be *invented*. This means that abduction can be more or less *creative*, which leads to classifications of various types of abduction (see the next section).

The answers to the two questions thus are as follows: (1) No, abduction (in some cases) cannot be based only on explicit knowledge because its *rule* is invented, thought-out, in the process and this seems to involve something like "tacit knowledge," i.e., something we know but cannot tell (see Polanyi, 2009); (2) The knowledge used in abduction also comes from abduction.

7 Types of Abduction

Massimo A. Bonfantini and Giampaolo Proni conclude their discussion of abduction in Holmes and Peirce by distinguishing three types of abduction, which are characterized by "ascending degrees of originality and creativity" with respect to the source of the "mediation law," i.e., in Peirce's terminology, the *rule*. It is:

(1) "given in an obligant and automatic or semiautomatic way"
(2) "found by selection in the available encyclopedia"
(3) "developed de novo, *invented*", or *guessed* (Bonfantini & Proni, 1983, pp. 133–134)

According to Bonfantini and Proni, Holmes' abductions are usually plausible and "sensible" (which we can see in the "Afghanistan" example above — after explanation, Watson has no problem accepting it) and *not very original*, and therefore more like "puzzle-solving" (p. 127). The important point is that they *must* be plausible, not too risky and, therefore, unoriginal because they are to be used in proving guilt (p. 128).[33] So, it seems that Holmes' abductions fall into the first two types, which is also supported by his claim that he *never guesses*.[34]

[32] See also Shead (2018, pp. 34–35).

[33] This, however, is not entirely true. Holmes does not just generate hypotheses; he also *tests* them (here, deduction and induction come into play as well). Therefore, even *prima facie* implausible hypothesis can become generally acceptable, just as in science.

[34] *The Sign of the Four,* p. 91. This claim, however, is not to be taken too seriously. See Sebeok and Umiker-Sebeok 1983, pp. 21–22. Cf. Holmes' admission of a "bluff" in *The Problem of Thor Bridge*, p. 1104.

Eco, in the same volume, follows this classification, calling the third type "creative abduction," and adding the concept of "*meta-abduction*," which must accompany the creative abductions and which

> consists in deciding as to whether the possible universe outlined by our first-level abductions is the same as the universe of our experience. (1983, pp. 206–207)

That means that if we creatively guess an explanation, we must do this additional "step" and take this explanation to be the true one.

Eco, however, sees many of Holmes' inferences as *creative* abductions and gives two examples (pp. 215–217). One of them is Holmes' inference from the fact that Watson has a "little reddish mould" on his instep to the conclusion that he had been to the post office on Wigmore Street and, then, from the fact that he did not write a letter during the morning and had a stockpile of stamps and postcards at home, that he sent a telegram from there. Eco analyzes the beginning of the reasoning as an abduction of the first type (in his terminology "overcoded"):

> people with mud adhering to their instep have been in an unpaved place, and so on

and one of the second type ("undercoded"): within the close neighborhood, only on Wigmore Street does earth have this particular tint, and it is more probable that Watson was there than somewhere farther away. But this is not enough to suppose that Watson, even if he happened to walk this specific street, visited the post office, especially when Holmes has no other reason to suppose that he went to any post office at all — he actually has reasons to think the opposite.[35]

The *second* inference takes it as a fact that Watson went to the post office and uses it as its premise. It is actually presented by Holmes as an application of the method of elimination (Holmes apparently cannot think of other reasons for going to a post office than those mentioned) and therefore, it actually does not seem as particularly *creative*. On the other hand, if we take it as a whole, inventing the *story*[36] that Watson went to a particular post office and sent a telegram based on just the fact that he had mud on his shoe and appeared to have *no* reason to go to a post office deserves to be called "creative" at least in the sense that it is by no means an abduction that we would generally make in a (semi)automatic manner or choose from several generally accepted possibilities. In other words, it is a little outlandish.

A more multifaceted classification of abductions is given by Schurz (2008). On a more abstract level, abductions can be classified as *selective* or *creative*, which agrees with the classifications mentioned above. Schurz's classification itself is constituted by various "*patterns*" of abduction that are classified using three dimensions

[35] But to be fair, Holmes probably thinks that one is especially likely to tread in the soil when entering the post office and not just walking down Wigmore Street: "Just opposite the Wigmore Street Office they have taken up pavement and thrown up some earth, which lies in such a way that it is difficult to avoid treading in it in entering" (*The Sign of the Four*, p. 89).

[36] Cf. Demeter (2011).

and hierarchically ordered (pp. 205–206). These patterns make explicit that there are abductions of significantly different kinds. The abductions Sherlock Holmes is most of the time concerned with would in this classification fall into *factual abductions* (more specifically, usually either *observable-fact abduction* or *first-order existential abduction*), in which from a fact we abduce another fact, according to a known law (see pp. 206–209).

In this classification, even law-abduction, i.e., abduction from known laws to a *new* law, is counted as ("mainly") *selective*. For example:

> All substances which contain molecular groups of the form C have property E.
> All substances of empirical kind S have certain empirical properties E.
> Conjecture: Substances of kind S have molecular characteristics C. (p. 212)[37]

This shifts the threshold of "creativity," excluding laws that could otherwise be counted as invented or *guessed* "de novo." *Creative* abductions, according to Schurz, are those that *introduce new concepts or models* (p. 202). These types of abduction are applied mainly in science and include *theoretical-model abduction* and *second-order existential abduction* with its subtypes (e.g., *micro-part abduction*, which postulates the existence of some micro-parts that constitute known objects and therefore introduces a new concept like "atoms").

For Peirce, abduction is the key to creativity in general,[38] but we have seen that it is a widely varied phenomenon,[39] that can be creative to a greater or lesser degree. Abductions of some kinds would not count as creative at all. Now, the questions for us are: Is Sherlock Holmes creative? and Could something like his "ideal reasoner" (i.e., basically a computer program) be creative?

I take Holmes' opinion on this matter to be that the "ideal reasoner" is *not* creative. It just "deduces" according to deterministic laws, so there is nothing to be creative about. If Holmes himself is creative in his "deductions," it is actually, from his point of view, an imperfection. But is he really creative? In the example of what is, according to Eco, a creative abduction, we have seen that taken as a whole, it seems like wild guessing. But there are several abductive steps involved that would not be counted as creative on their own. So, if abductions like those should be called creative, perhaps creativity consists in combining more "overcoded" and "undercoded" abductions in a wholly new way and often *against all odds*.

This looks on the one hand, as if creativity could be reduced to "*combinatorics*" and, therefore, something a computer could easily do. On the other hand, however, there could obviously be myriads of alternative combinations of simple abductions arising from the same facts. Generating the combinations is not enough to come up with a creative solution, then. Choosing just one of the possibilities is an equally important part of the process. But how to choose just one? One has to have a method

[37] Eco (1983) also distinguishes between abductions *from facts to facts* and from *facts to laws*.

[38] See for example Poltronieri (2019, p. 303).

[39] According to Peirce, it is even wider than these classifications suggest because even *sensation* and *emotion* are related in an important way to "hypothesis" (see Peirce, 1868, pp. 149–150; 1878, p. 482).

for selecting the *best* explanation from the possibilities, hence the conception of abduction as *the inference to the best explanation*.

Here we have the problem of completeness of knowledge again: we normally cannot be sure that the possibilities that we can generate include the right one. Accordingly, we are actually limited to inference to the best *available* explanation (see Schurz, 2008, p. 203). But an even bigger problem is that the notion of the *best* must be given some clear sense. The "ideal reasoner" must somehow *calculate* which explanation is the best. It must, for example, as Holmes says, "balance probabilities."[40] What is more likely, that Watson went to the post office given that he didn't need anything there and had no letter to send, or that he got his shoes dirty somewhere else than on Wigmore Street? In every case, the reasoner would have to have the probabilities associated with the "laws" in the knowledge base. Where to get those and how to make sure that their utilization will lead to the correct conclusions in novel cases? The "best explanation" can very easily go wrong, unless the ideal reasoner knows so much that there is no space for alternatives. So, on the one hand, the ideal reasoner would have to generate many hypotheses and evaluate them in an exact manner to mimic Holmes's creativity (and his "meta-abduction"). But on the other hand, Holmes might think that he has to be creative to *compensate* for his lack of knowledge.

8 Logic of Discovery and Creativity

The notion of abduction is related to a problem discussed in the philosophy of science in the last several decades: Can there be a *logic of discovery*? (Burks, 1946, p. 302). Because if so, abduction seems like a suitable candidate for something that can bridge between *logic* on one side (as a kind of *inference*) and *discovery* on the other (as something that can bring *something new*). The opposition between *discovery* and *justification* (in which logic plays a role without doubt) is also reflected in different accounts of abduction, one of them being the *inference to the best explanation*, which includes the justification side: we infer that particular explanation is true *because* it is the best one. Another option is to take abduction as just generating hypotheses (cf. Magnani, 2001, pp. 19, 78). In that case, it should be seen as one phase or task (Levi, 2006, pp. 281–282) of scientific work, the others being deduction (of consequences of the hypotheses) and induction (testing of the hypotheses). This is how Peirce sees abduction later (Peirce, 1974, 6.469ff, 7.202ff.). The distinction between justification and discovery also corresponds Schurz's distinction between *justificational* functions and *strategical* functions of inference. Schurz takes different kinds of inference to have these functions to a different degree, deduction having *mainly* a justificational function, abduction mainly (but not solely) a strategical function (Schurz, 2008, pp. 203–204).

The possibility of a "logic of discovery" is famously dismissed by Karl Popper (see Simon, 1973, p. 471). N. R. Hanson, on the other hand, criticizes both inductive and hypothetico-deductive ("H–D") accounts of (physical) theory, the latter because it misrepresents what physicists actually do — they do not start from hypotheses

[40] *The Hound of the Baskervilles*, p. 516.

but from *data*. That is what, according to Hanson, got the inductive account right. But the hypotheses, or laws, are not induced; they are actually not just generalizations of the data. They are *explanations* (Hanson, 1958, p. 70f). That is why Hanson draws on Peirce's conception of abduction (his preferred term is "*retroduction*") and claims that discovery is a matter of logic:

> Disciples of the H-D account often dismiss the dawning of an hypothesis as being of psychological interest only, or else claim it to be the province solely of genius and not logic. They are wrong. If establishing an hypothesis through its predictions has a logic, so has the conceiving of an hypothesis. To form the idea of acceleration or of universal gravitation does require genius: nothing less than a Galileo or a Newton. But that cannot mean that the reflexions leading to these ideas are unreasonable or a-reasonable. (Hanson, 1958, pp. 71–72)

And, as it seems, Hanson finds conceiving of a hypothesis much more interesting than the deductive and testing parts of scientific work, i.e., those parts that according to Popper belong to the province of logic.

Hanson characterizes retroduction as "[p]erceiving the pattern in phenomena" (p. 87). Simon, drawing on Hanson's work, wants to specify the "logic of discovery" in more concrete terms. The "logic of scientific method" consists of *normative* statements about what to do (what "processes" to employ) to reach a given *goal*, i.e., discovery of valid scientific laws, just as the "'logic' of a chess strategy" would be a set of normative statements telling us what to do to checkmate the opponent's king (Simon, 1973, p. 473). So, there can be a logic to finding the right "pattern" in the data (like the shape of the orbit of Mars in the observations). Such a "logic" can be very simple, like "check all possibilities and see which one fits."

The processes prescribed by such *logics* can be, of course, more complicated, and also more *efficient*. This is the case of *heuristic* search algorithms (Simon, 1973, p. 476; Zytkow & Simon, 1988; see also Genot, 2018, p. 2080), which work in a "smarter" way by reducing the space of possibilities that must be tried. This is heralded already by Peirce's theory of abduction: Because there are a potentially unlimited number of possible hypotheses, part of the problem of abduction is the problem of the *economy* of research (Fann, 1970, pp. 47–51), which is basically the problem of selecting which hypothesis to try first.

So, for example, Holmes' ideal reasoner would, given some facts from the crime scene, generate many hypotheses explaining what happened, say a couple of thousands, each composed of a "chain" of possible facts like "The murderer got away through the window." A heuristic is some rule (or a "trick") that allows reducing this *space* of possible explanations to a manageable size so that they can be "tested." Holmes apparently can do this, for example, by not even thinking about hypotheses that would be too unlikely, for example, if they include some supernatural explanation. A simple example of a heuristic that could be used by a detective, human or computer, could be based on something like "The murderer usually knows the victim." This tells us which of the possible explanations to consider first and which to push aside as less probable.

An example of a heuristic utilized in a system of discovery of scientific laws that looks for linear relationships between variables can be testing if one variable changes monotonously with the growth of another variable (that is not linearly related to it) and if so, adding their ratio as another variable (Zytkow & Simon, 1988, p. 70).

Heuristics of this kind are implemented in great many AI systems, including "normative systems of discovery" that aim to discover scientific laws (see, e.g., Zytkow & Simon, 1988; Magnani, 2001, pp. 49–52). So, it seems that not only is there a possibility of *a* logic of discovery, but there *are* such logics that are explicit enough — in the form of algorithms and computer programs.[41] Now, the proponents of a logic of discovery can see this as showing that it is possible to capture the *creative* process algorithmically, or, in other words, that being creative amounts to utilizing the right heuristics. In the end, if we can algorithmically specify the right tricks for finding valid scientific laws, then why could we not specify tricks for creating, for example, novel plots or musical themes as well.

But there is a problem with taking such systems as *a model of creativity* in general. Let us admit that new ideas are usually a combination of already known elements and that some ways to combine those elements (or to *find* the right combination) that are oftentimes successful can be captured by heuristics. That is why there are so many relatively successful *specialized* systems for, for example, generating pictures, music, or texts. But if these systems use heuristics, *someone had to invent these very heuristics*. Could they be invented using some other heuristics? In principle, sure. But I do not see how the most general notion of creativity could be captured in this way, without creating an infinite regress or postulating some general formula of creativity, a heuristic prescribing how to invent anything at all, including this heuristic itself.

There are also, of course, modern approaches to AI that are based on artificial neural networks and differ in important respects from the systems of classical symbolic AI. Those systems include, for example, language models like recently very popular GPT-3 (Brown et al., 2020), on which the chatbot ChatGPT is based. However, such systems and the possibility of their creativity are out of scope of this article, as their connection with abduction and with Holmes' theory is not so straightforward.[42]

[41] Cf. Lorenzo Magnani: "The [epistemological] theories [of reasoning] and the programs are, quite literally, two different ways of expressing the same thing" (2001, pp. 51–52).

[42] Artificial neural networks are *trained* on large amounts of data, if we put it in an abstract and simplified way, to create their own *generalizations* from the data that can be applied to new data. For example, a neural network trained to recognize cars, traffic lights etc. in pictures can recognize these kinds of objects, because it has some generalized representations of them encoded in the trained *weights* of the connections between its "neurons." In the case of language models, the training data are huge corpora of texts, and the network is trained (again put very simply) to continue from a given prompt, i.e., generate text that would be likely to follow, given the short input and the knowledge how the texts used for training look like.

These trained generalizations can be seen as implicit knowledge, which has a role somehow analogous to the role of heuristics in systems of the symbolic AI paradigm discussed here. The important difference is that, unlike the heuristics, this "knowledge" was not explicitly coded into the system but was created in the course of training by the algorithm (in something like an *abductive* inference, we could say, because it starts from *data* and leads to some "general rule" – although, in this case, usually not an explicit one – or, in Hanson's words, it is some "pattern" that can be seen in the "phenomena," i.e., data). That is why the paradigm of neural networks seems to be much more promising for modeling general creativity and intelligence. But the matter is, of course, not so simple, and certainly cannot be tackled here.

9 Conclusion

Sherlock Holmes with his "ideal reasoner" can be seen as an ideologist of symbolic AI. He sees the ideal detective (or at least his "reasoning part") as a computer that not only mimics the reasoning of a real detective but makes it more efficient and accurate thanks to better knowledge and exact reasoning processes. Creativity should not play a role here, which is actually part of the point and is reflected in Holmes' contention that the reasoning involved could be characterized as deduction.

There are, however, notorious problems with this idea, which brings *abduction* into play. Abduction is not only a better characterization of Holmes' actual reasoning, but also a much better guide for designing "reasoning machines." Abduction (as understood by Peirce) comes hand in hand with creativity. Not every kind of abduction deserves to be called "creative," but a germ of creativity can be seen already in selecting from multiple possibilities (while both generating them, if they are not simply predefined, and choosing the most "likely" may belong to the creative process). This, on the one hand, undermines Holmes' idea again, to some degree (the computer detective would not be as exact as he might have hoped); on the other hand, it introduces the idea of the possibility of creative machines.

The creativity of such machines could be explicable by something like a "logic of discovery." This seems to be viable in specific cases, but such a logic also has to be "discovered" (or invented), presumably by (*creative*) abduction. Creativity can be seen as a matter of degree, which corresponds to the various kinds of abduction. Some tasks which would be considered creative can arguably be done by AI, but whether it will be someday capable of creativity in the most general sense remains to be seen.

Funding Open access publishing supported by the National Technical Library in Prague.

Declarations

Conflict of Interest The author declares no competing interests.

References

Bonfantini, M. A., & Proni, G. (1983).To guess or not to guess? In Eco & Sebeok (1983), 119–134.

Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., … Amodei, D. (2020). *Language models are few-shot learners*. https://doi.org/10.48550/ARXIV.2005.14165

Burks, A. W. (1946). Peirce's theory of abduction. *Philosophy of Science, 13*(4), 301–306. https://doi.org/10.1086/286904

Demeter, T. (2011). A touch of the dramatic. In J. Steiff (Ed.), *Sherlock Holmes and philosophy: The footprints of a gigantic mind*. Open Court.
Doyle, A. C. (2013). The complete Sherlock Holmes. Race Point Publishing.
Dreyfus, H. L. (1992). *What computers still can't do: A critique of artificial reason*. MIT Press.
Eco, U., & Sebeok, T. A. (Eds.). (1983). *The sign of three: Dupin, Holmes.* Indiana University Press.
Eco, U. (1983). Horns, hooves, insteps: Some hypotheses on three types of abduction. In Eco & Sebeok (1983), 198–220.
Fan, F.-L., Xiong, J., Li, M., & Wang, G. (2021). On interpretability of artificial neural networks: A survey. *IEEE Transactions on Radiation and Plasma Medical Sciences, 5*(6), 741–760. https://doi.org/10.1109/TRPMS.2021.3066428
Fann, K. T. (1970). Peirce's theory of abduction. *Martinus Nijhoff*. https://doi.org/10.1007/978-94-010-3163-9
Genot, E. J. (2018). Strategies of inquiry: The 'Sherlock Holmes sense of deduction' revisited. *Synthese, 195*(5), 2065–2088. https://doi.org/10.1007/s11229-017-1319-x
Genot, E. J. (2020). The Holmesian logician: Sherlock Holmes' "science of deduction and analysis" and the logic of discovery. *Synthese*. https://doi.org/10.1007/s11229-020-02709-w
Grice, H. P. (1975). Logic and conversation. In Cole, P., Morgan, J. (Eds.), *Syntax and semantics: Vol. Vol. 3* (pp. 41–58). Academic Press.
Hanson, N. R. (1958). *Patterns of discovery: An inquiry into the conceptual foundations of science*. Cambridge University Press.
Harrowitz, N. (1983). The body of the detective model: Charles S. Peirce and Edgar Allan Poe. In Eco & Sebeok (1983), 179–197.
Hintikka, J., & Hintikka, M. B. (1983). Toward a theory of information-seeking through questioning. In Eco & Sebeok (1983), 154–169.
Hobbs, J. R., Stickel, M., Appelt, D., & Martin, P. (1993). Interpretation as abduction. *Artificial Intelligence, 63*(1–2), 69–142. https://doi.org/10.1016/0004-3702(93)90015-4
Klinefelter, W. (1989). The writings of Mr. Sherlock Holmes. In Shreffler, P. A. (Ed.), *Sherlock Holmes by gas-lamp: Highlights from the first four decades of the Baker Street Journal*. Fordham University Press.
Levi, I. (2006). Beware of syllogism: Statistical reasoning and conjecturing according to Peirce. In C. J. Misak (Ed.), *The Cambridge companion to Peirce* (pp. 257–286). Cambridge University Press.
Magnani, L. (2001). *Abduction, reason, and science: Processes of discovery and explanation*. Kluwer Academic/Plenum Publishers.
Peirce, C. S. (1868). Some consequences of four incapacities. *The Journal of Speculative Philosophy, 2*(3), 140–157.
Peirce, C. S. (1878). Deduction, induction, and hypothesis. *Popular Science Monthly, 13*, 470–482.
Peirce, C. S. (1974). In C. Hartsthorne, & P. Weiss (Eds.), *Collected papers of Charles Sanders Peirce*. Belknap Press.
Pigliucci, M. (2012). Sherlock's reasoning toolbox. In P. Tallon & D. Baggett, *The philosophy of Sherlock Holmes*. University Press of Kentucky.
Polanyi, M. (2009). *The tacit dimension*. The University of Chicago Press.
Poltronieri, F. A. (2019). The scientific enterprise illustrated: Abduction, discovery and creativity. In A. Görgen, G. A. Nunez, & H. Fangerau (Eds.), *Handbook of popular culture and biomedicine: Knowledge in the life sciences as cultural artefact* (pp. 301–306). Springer.
Schurz, G. (2008). Patterns of abduction. *Synthese, 164*, 201–234. https://doi.org/10.1007/s11229-007-9223-4
Sebeok, T. A., & Umiker-Sebeok, J. (1983). "You know my method": A juxtaposition of Charles S. Peirce and Sherlock Holmes. In Eco & Sebeok (1983), 11-54.
Shead, J. (2018). Do we know his methods? Ratiocination in the works of Arthur Conan Doyle. *Clues, 36*(1), 20–40.
Simon, H. A. (1973). Does scientific discovery have a logic? *Philosophy of Science, 40*(4), 471–480. https://doi.org/10.1086/288559
Smajić, S. (2010). Ghost-seers, detectives, and spiritualists: Theories of vision in victorian literature and science. *Cambridge University Press*. https://doi.org/10.1017/CBO9780511712012
Zytkow, J. M., & Simon, H. A. (1988). Normative systems of discovery and logic of search. *Synthese, 74*(1), 65–90. https://doi.org/10.1007/BF00869619