Rationality in discovery

Bosch, Alexander Petrus Maria van den

Part II Discovery

What is the rational use of theory and experiment in the process of scientific discovery, in theory? In this part I discuss three different approaches to the study of the rational use of theory and experiment in the process of scientific discovery. I start with a discussion of the study of **logic** (Chapter 4). Then I discuss an account that stems from the psychological study of **cognition** (Chapter 5). I finish this part with the discussion of a model of discovery that is grounded in the study of **computation** (Chapter 6).

Chapter 4

Logic

4.1 Introduction

In this thesis we set the general problem: what is rationality in scientific discovery? This question receives attention from several academic disciplines. Traditional philosophers of science are usually interested in what scientific discovery ought to be, and how reasoning in that process can be valid or justified. Empirical scientists are usually more interested in describing rationality in scientific discovery as a social or psychological phenomenon, to be studied empirically.

In this chapter we will address a normative approach that stems from studies in logic. In the next chapter we will address a psychological theory about the rationality of reasoning and problem solving. This part will end with a chapter on a general computational model of discovery. In discussing all models I will look for answers to the specific questions from section 1.3, *i.e.* those about: (1) the structure of a theory, (2) the process of scientific reasoning and (3) the route between theory and experiment.

In this chapter we start with a discussion of logic, the traditional study of valid reasoning. The question is: what is the rational use of theory and experiment in the process of scientific discovery, as proposed in the study of logic? We start by asking: what is a scientific theory and what is scientific reasoning?

To address these questions I discuss an illustrated example of explanation. In an episode of the life and times of cartoon character Calvin and his tiger Hobbes he watches a sunset with his father, see Figure 4.1. His father explains the setting of the sun to Calvin. Now, why would we not accept his explanation as scientific? Is this because his hypothesis is not scientific? Is this because his reasoning is not valid? Let us look at the validity of his inferences from the perspective of logical argumentation theory, and reconstruct his reasoning.

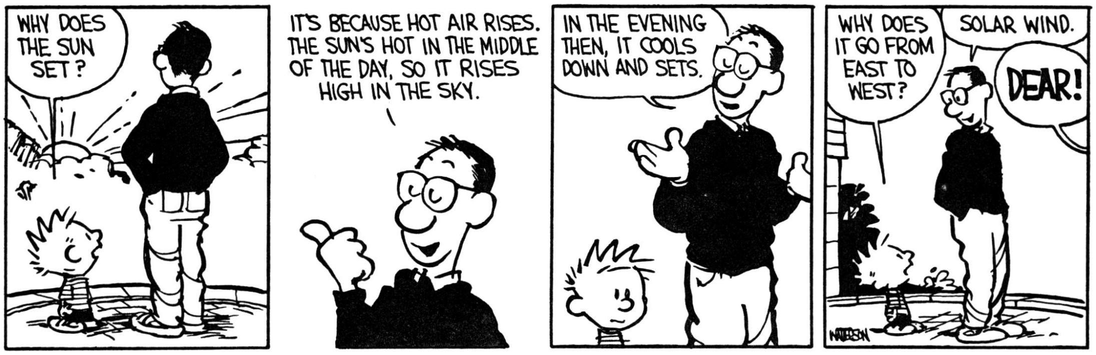

Figure 4.1: Calvin and Hobbes © 1988 Bill Watterson. Reprinted by permission of Universal Press Syndicate. All rights reserved.

4.2 Deduction

Calvin's question is: why does the sun set? This question asks for an explanation of his observation. He wants to know what causes the sun to set. If Calvin accepts only a logically valid answer, he can only accept as explanation a deduction of his observation from what is known. Let us examine his inferences one at a time and comment on their validity.

In modern logic the validity of an inference is independent of the truth of the premises. Yet when an inference kind is valid the conclusion is true when the premises are true. To represent kinds of inference schemes in the discussion I will use a two or three letter abbreviation (TLA) that is *italicized* if it represents a logically valid inference. In an inference scheme I will mark a proposition with a star (*) to indicate that we do not know whether that proposition is true.

Calvin's father manages to infer his answer in several possibly implicit steps. First he presupposes two propositional premises as initial assumptions which Calvin should accept off-hand, without further argumentation:

P_1 Hot air rises	Hot(air) ⇒ Rises(air)
P_2 In the middle of the day the sun is hot	Hot(sun)

Presumably he further assumes that the air is hot, and that the sun causes it:

P_3 If the sun is hot then the air is hot	Hot(sun) ⇒ Hot(air)

These premises seem unproblematic. Based on them he can validly infer by *modus ponens* (*MP*) that the air is hot:

P_2 In the middle of the day the sun is hot	Hot(sun)	
P_3 If the sun is hot then the air is hot	Hot(sun) ⇒ Hot(air)	
———	———	*MP*
P_4 In the middle of the day the air is hot.	Hot(air)	

By transitivity (*TRN*) he can infer validly that if the sun is hot the air rises:

P_1 Hot air rises	Hot(air) ⇒ Rises(air)	
P_3 If the sun is hot then the air is hot	Hot(sun) ⇒ Hot(air)	
		TRN
P_5 If the sun is hot then the air rises.	Hot(sun) ⇒ Rises(air)	

From the premises he then infers in two steps why the sun rises. To be explained first is the observation:

P_6 In the middle of the day the sun rises Rises(sun)

This should be a conclusion from our premises and valid intermediate conclusions. For the first step three different inferences are possible. The first could be:

P_5 If the sun is hot then the air rises.	Hot(sun) ⇒ Rises(air)	
		GEN
P_7 If the sun is hot, anything rises *	for all x Hot(sun) ⇒ Rises(x) *	

Logically this is a fallacy, a hasty generalization (GEN) called a *secundum quid*. Seeing one type of object with a property does not imply that all have the same property. So this inference is invalid. However, we can not say that, logically, his conclusion is false either. The inferred proposition could well be true, but its truth does not follow deductively from the truth of the premise.

Alternatively, Calvin's father could have assumed that the sun is part of the air and a property of air is also a property of the sun. Since hot air rises, a hot sun rises as well:

P_1 Hot air rises	Hot(air) ⇒ Rises(air)	
P_8 The sun is part of the air	sun part of air	
		DVS
P_9 If the air is hot the sun rises *	Hot(air) ⇒ Rises(sun) *	

This is known as a *fallacy of division* (DVS), where a property of the whole is also ascribed to a part. All the parts together could well not have the same property as the whole (*e.g.* the parts are light, but the whole is heavy). Yet again, it is also possible that the whole does have the same property as the parts, and vice versa (*e.g.* the whole is light, therefore each part is light).

A third possible interpretation of the explanation of Calvin's father is a *causal argumentation* (CAU):

P_4 In the middle of the day the air is hot.	Hot(air)	
P_6 In the middle of the day the sun rises	Rises(sun)	
		CAU
P_9 If the air is hot the sun rises *	Hot(air) ⇒ Rises(sun) *	

In this case a cause-effect relation is inferred from the mere observation that two events take place together. The air is hot and the sun rises, hence if air is hot the sun will rise. This is again a logical fallacy. The causal relation could just as well be the other way around, or not existent. The occurrence of events one after another could just as well be an incident. This fallacy is called *post hoc ergo propter hoc*.

We now saw three ways to infer P_9 in a first step. To explain P_6, the rising of the sun, he further infers in the second step:

P_9 If the air is hot the sun rises *	Hot(air) ⇒ Rises(sun) *	
P_4 In the middle of the day the air is hot.	Hot(air)	
		AA
P_6 In the middle of the day the sun rises *	Rises(sun) *	

In this inference the second premise affirms the antecedent (*AA*) of the first premise. This inference is called *modus ponens*. It is a valid inference that guarantees the truth of the conclusion if the premises are both true. But in this case the conclusion may be false because the first premise may not be true. So P_6 follows validly from P_4 and P_9, but not from our initial premises P_1 to P_4, because P_9 does not follow from them.

But Calvin's question was why the sun *sets*. To explain this, his father first implies in a third step that when the sun is not rising the air is also not hot.

P_9 If the air is hot the sun rises *	Hot(air) ⇒ Rises(sun) *	
P_{10} In the evening the sun sets.	not Rises(sun)	
		DC
P_{11} In the evening the air cools down. *	not Hot(air) *	

The second premise denies the consequent (*DC*) of the first. This inference is called *modus tollens*. Just like in an affirmation of the antecedent, the conclusion of the inference is true if the premises are true. We cannot say that for the first premise, so the conclusion may not be true.

To conclude the explanation his father further treats a possibly sufficient condition as a necessary condition. From the assumption that the sun rises when the air is hot he infers that when it is not hot, the sun also does not rise, thereby Denying the Antecedent (DA) of the first premise. This is also called an inverted *modus tollens*. This inference is invalid. Hence, the conclusion may be false even when the premises are all true.

P_9 If the air is hot the sun rises *	Hot(air) ⇒ Rises(sun) *	
P_{11} In the evening the air cools down.	not Hot(air)	
		DA
P_{10} In the evening the sun sets *	not Rises(sun) *	

Assuming he translates not hot air (P_{11}) with cool air (P_{13}) and a not rising sun (P_{10}) with a setting sun (P_{14}), he rephrases this conclusion (invalidly) in the statement with P_{13} and P_{14} as premises: if the air cools down the sun sets (P_{12}), which given that the air cools would validly imply that the sun sets if the statement were true:

P_{12} If the air cools down the sun sets *	Cools(air) $\Rightarrow$ Sets(sun) *	
P_{13} In the evening the air cools down.	Cools(air)	
		AA
P_{14} In the evening the sun sets *	Sets(sun) *	

We can also interpret his whole explanation in yet another way. He could have assumed that the premise: if the sun rises the air is hot, stated that the rising sun is a necessary condition for hot air and hence infer that if the sun sets the air cools, via contraposition. But then he Affirms the Consequent (*AC*) of this proposition, also called an inverted *modus ponens*, to infer that if the air cools, the sun sets:

P_{15} If the sun sets the air cools down *	Sets(sun) $\Rightarrow$ Cools(air) *	
P_{13} In the evening the air cools down	Cools(air)	
		AC
P_{14} In the evening the sun sets *	Sets(sun) *	

By no great surprise this is invalid since the antecedent is not a necessary condition but a sufficient condition. In that case when the consequent of the first premise is known to be true the antecedent could be true, but could possibly be false just as well.

What can we conclude from this? Today, Calvin's father's explanation is gathered to be wrong. But is this because his hypothesis is unscientific, or because many of his inferences are fallacies? If we look at the beginning of modern science, three centuries ago, then what would we expect?

The Inquisition

In the seventeenth century Galileo Galilei defended the Copernican heliocentric theory. This theory put the sun at the center of the solar system, and explained that the sun sets because the earth turns on it own axis and revolves around the sun. It also explained the phases of Venus that Galileo first observed with his self made telescope.

Venus waxes and wanes as viewed from the earth, similar to the moon's phases. When Venus is full, we cannot see it because the sun is in the way. As Venus wanes from the full phase, it also gets bigger because it is approaching us. When it is closest to us, we cannot see it because no light is reflected towards us. This could be explained if it was assumed that both Venus and the Earth rotate around the sun. If you put the earth in the center then you could only explain it when you assumed Venus to rotate around the sun while Venus and the sun both rotate around the earth.

In 1616 Galileo was formally warned by the church to stop this defense. The reason for this censure was not that the claim was considered wrong, or that teaching so undermined the Church. Rather, it was claimed that Galileo's proof for the theory was not logically valid. Galileo's main argument depended on the fact that the theory explained why the planet Venus shows phases. Yet, he could not prove this deductively. The argument ran as follows:

If the planetary system is heliocentric, then Venus will show phases.
Venus shows phases.

Hence, the planetary system is heliocentric

So the argument was based on an affirmation of the consequent, a fallacy well known by the Aristotelian clergy. While Venus does indeed show phases, the planetary system being heliocentric may not be the only condition under which that is true. The clergy pointed out the flaw and Galileo was ordered not to put forth this idea as proved.

Pope Urban VIII, who just as Galileo was a member of the Academy of Lynxes, a scientific society formed in 1603, informally lifted these orders in 1633. There is evidence that the Pope gave Galileo the opportunity to neutrally compare the heliocentric theory with the geocentric system of Ptolemy, and come up with a deductive proof.

But in the book he then wrote he patronized the Pope, who was greatly offended. As a result Galileo was accused of disobeying the order of 1616 to stop his defense of the Copernican system. Even though Galileo could produce a letter that showed he was merely warned instead of ordered, he was threatened by the Inquisition, shown the “instruments” (of torture), and sentenced to house arrest for the rest of his life. Hence, it was disobeying orders to stop using a fallacy that got him convicted by the Inquisition, and not committing heresy, since technically the Copernican system was never declared heretical (Gingerich, 1992). However, today science accepts Galileo’s explanation. But is this because his reasoning is scientific?

4.3 Induction

A typical scientific explanation can never deductively follow from what we already know or have observed, because most scientific hypotheses include assumptions and predictions about future or other not observed situations. It is logically always possible that those situations will be different.

In his defense of the Copernican system Galileo not only needed to defend a scientific theory, but also a manner of reasoning. Galileo employed an inductive inference. The conclusion of an inductive inference can contain more or other information than its premises, hence it is not deductively valid. Deductive inference preserves the truth of its premises so as to encompass its conclusion, while an inductive inference expands beyond them.

However, scientific reasoning is not void of deductive reasoning. Logicians consider a sound explanation to be a deductive conclusion from a number of true hypotheses. The problem with scientific hypotheses is that you can never know for sure whether they are true. The philosopher Karl Popper stressed that what you validly can know about a hypothesis is that it is false. If a hypothesis claims that all particulars of a type have a property, then only one particular of the type without that property will validly imply that the hypothesis is incorrect.

At the beginning of this century, the philosopher Charles Sanders Peirce coined the term ‘abductive inference’ to distinguish Galileo’s inference from other kinds of inductive inference like generalization (GEN). With generalization you infer that if a number of particulars of a type have a property, then all particulars of that type have that property. So for example:

The fact that a number of particulars of type A have property C is observed;

Hence, there is a reason to suspect that all A have property C

According to Peirce the function of abduction is *ampliative*, to introduce new ideas. A hypothesis suggested by abduction should contain predictions about other properties or other types of particulars as well. In his later work Peirce (1958, 5.188) put forward the following often quoted definition of abductive inference:

"The surprising fact, C, is observed;
But if A were true, C would be a matter of course.

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

Abductive inference is actually part and parcel of everyday common sense reasoning. But it seems that it can lead to the wildest of explanations, as Calvin can attest. But even though his father commits enough deductive fallacies to experience more of the "instruments" then just their sight, had he lived three centuries ago, his explanation is not problematic just because of its inductive nature. Both Galileo and Calvin's father seem to follow the same inference. But then what makes Galileo's inference differ from that of Calvin's father's?

4.4 Abduction

How does Peirce's definition of abduction compare to other kinds of inductive inferences? Let us take a closer look at Peirce's inference scheme and our examples. We will address the similarities and differences. The examples are summarized in Table 4.1. The properties Hot and Rises are abbreviated to H and R respectively.

Inference	Premise 1 (Observed C)	Premise 2 (If A where true C follows, if)	Conclusion (A)
GEN	H(sun) ⇒ R(air)	∅	∀x H(sun)⇒R(x) *
DVS	H(air) ⇒ R(air)	sun part of air	H(air) ⇒ R(sun) *
CAU	R(sun)	H(air)	H(air) ⇒ R(sun) *
AA	H(air)	H(air) ⇒ R(sun) *	R(sun) *
DC	Not R(sun)	H(air) ⇒ R(sun) *	Not H(air) *
DA	Not H(air)	H(air) ⇒ R(sun) *	Not R(sun) *
AC	Cools(air)	Sets(sun) ⇒ Cools(air) *	Sets(sun) *
AC	Phases(Venus)	Center(sun)⇒ Phases(Venus)	Center(sun) *

Table 4.1: Summary of examples of the discussed inference types

For an inference to fit Peirce's definition of abduction, Premise 1 should follow as a matter of course if Premise 2 and the Conclusion are both assumed to be true. If we compare the example inferences with this definition we notice that generalization (GEN), division (DVS) and causality (CAU) fit the definition well. If the conclusion and Premise 2 are true, then premise 1 is also true. If a generalization is true, then the truth of a particular follows as a matter of course. In the division example Premise 1 follows based on premise 2 and the assumption in the conclusion that the property of a whole is also a property of its parts. If the causal implication in a conclusion and premise 2 would be true, premise 1 would follow by *modus ponens*. In sum, the inferences GEN, CMP and CAU can be seen as special kinds of abduction, according to Peirce's definition.

The next two types in Table 4.1 do not fit the definition. Not remarkably these are the deductively valid inferences affirmation of the antecedent (*AA*) and denial of the consequent (*DC*). These will of course not fit a definition of an abductive inference. In abduction the conclusion is an explanation, in deduction the premises are. However, in the example Calvin's father used these inference kinds incorrectly, because he wrongly assumed the premises were true, to conclude the truth of the conclusion. Denial of the antecedent (*DA*) fits the definition well. The implication in Premise 2: if H(air) then R(sun) is logically equivalent to: if not R(sun) then not H(air). Not surprisingly the affirmation of the consequence (*AC*) most resembles the definition of abduction. The observed fact C affirms the consequent of $A \Rightarrow C$, where A is the conclusion. Both the explanation of the sun set and the phases of Venus follow that inference.

However, there is an important difference between the two. Premise 2, the implication if A then C, is true in the case of Galileo but uncertain in the case of Calvin's father. The implications are actually of a different nature. One is itself a hypothesis and the other a logical consequence. The former consists of a so called material implication and the latter of a logical or semantic implication. Let us take a closer look at the nature of implication and its role in Peirce's definition.

Implications

A material implication is a conditional statement that connects two independent statements. These statements may be either true or false depending on other conditions. The material implication asserts that when the antecedent is true, the consequent will also be true. Because of this property of the conditional statement it is argued that the material implication can represent a causal relation between two events described by the antecedent and the consequent. However, the truth of the conditional statement can already be settled by the status of only one of its constituents. The material implication is by definition already true when the antecedent is false or the consequent is true. Let us look at Table 4.2 to follow this.

Antecedent	Consequent	$A \rightarrow C$
True	True	True
False	False	True
False	True	True
True	False	False

Table 4.2: Truth table of the material implication

In Table 4.2 I summarize all possible truth value combinations of a material implication. The material implication is false only if the antecedent is true and the consequent is false. The statement "if the air cools down the sun sets" is such a statement. It states that it will not happen that the sun does not set while the air does cool down (row 4). It still allows for the possibility that the sun sets even though the air does not cool down (row 3).

Now let us look at Galileo's statement. If the planetary system is heliocentric, then Venus will show phases. This implication differs in nature. The antecedent statement logically implies the consequent statement, and many others. For instance it will also imply under what conditions the sun will set. When you say that the antecedent is true, you say that all its consequences are true, by implication. Formally we say that all models that make the propositions true that make up the antecedent of a semantic implication, will also make the consequent true. The models of the antecedent constitute a subset of the models of the consequent. As a notation we will, following tradition, use $A \models C$ for semantic implication, and $A \rightarrow C$ will denote material implication. For the language of predicate logic it has been proved that if C is semantically implied by A, it can also be deduced from A, written as $A \vdash C$, and vice versa.

Another important difference between material and semantic implication is shown by the set of the inferences that each allows. Given that C is true, it can be inferred that $A \rightarrow C$ is true, but you cannot infer the truth of A. Yet given that C is true, it *cannot* be inferred that $A \models C$ is true, but you can say that A is confirmed. However, this is only the case if $A \models C$ is true. Even if you assume that $A \rightarrow C$ is true, but $A \models C$ is not, then C does not confirm A. If it is known that $A \models C$ is true, you can (non deductively) infer the truth of A if all its consequences are confirmed.

Let us return to the definition of abduction. Apparently Peirce meant "if A were true then C would be a matter of course" to be a semantic implication. Abduction based on a semantic implication will introduce a hypothesis that may have many other implications. Hence the use of the term abduction: it forces alien statements into the explanation.

A generalized material implication, such as $\forall xy(A(x) \rightarrow C(y))$, may also entail new predictions, but they are usually about the same properties, A and C, and the same kind of objects, all x and y such that $A(x) \rightarrow C(y)$. The antecedent of a semantic implication, such as $A \models C$ may entail predictions about different properties and objects as well.

Galileo's inference was based on a semantic implication and Calvin's father assumed a material implication. But are abductions based on material implications unscientific? If that where so then we could not use laws to explain phenomena. The material implication "if the atmosphere pressure drops the air will cool down" could then not be used to explain why the weather cools down. Even Galileo's explanation of the phases of Venus would run into trouble. His hypothesis entails many material implications as possible consequences, *e.g.* (using abbreviations):

C: {position i of the sun and Earth → phase j of Venus}
A: {Center(sun)} ⊨ C: {position i of the sun and Earth → phase j of Venus}

A: {Center(sun) *}

To explain a particular phase of Venus an abduction could infer a particular position of planets. That would not necessarily need a semantic implication to be scientifically acceptable. A law could be formulated that relates the position of the sun and Earth to the phases of Venus, that could explain a particular phase on the basis of a particular position:

$$\frac{\begin{array}{l}\{\text{C: phase i of Venus}\}\\ \{\text{A: position x of the sun and Earth} \rightarrow \text{C: phase y of Venus}\}\end{array}}{\{\text{A: position j of the sun and Earth *}\}}$$

In many scientific areas not much more is known than material laws. So it may be desirable for Peirce to infer a rich logical hypothesis, but a material implication is not unscientific by its nature.

Definitions

The main difference between affirming the consequence of a material implication and affirming the consequence of a semantic implication is a difference in category. The former is part of the latter. Let us call the former kind *material abduction* and the latter kind *semantic abduction*:

$$\frac{\begin{array}{l}C\\ A \rightarrow C\end{array}}{A}\ \text{material abduction} \qquad \frac{\begin{array}{l}C\\ A \models C\end{array}}{A}\ \text{semantic abduction}$$

To avoid confusion between the two I will adopt the following notation. I will use the propositions C, A → C, and A, etc. to talk about statements that a semantic abduction reasons about. The premises and conclusion of a semantic abduction are sets that contain these statements. The first will be a set called P, containing a proposition about the world; the second premise a set H containing the hypothesis statement(s) that together with background assumptions B implies P. I can now define the different kinds of abduction as follows:

Definition 1 *Semantic abduction.* A semantic abduction is an inference that affirms the consequent of a semantic implication (ACS). Given the antecedent B ∪ H that semantically implies P, the affirmation of the consequent P infers hypothesis H:

$$\frac{\begin{array}{l}\text{Proposition P}\\ \text{Background B} \cup \text{Hypothesis H} \models \text{Proposition P}\end{array}}{\text{Hypothesis H: } \{*\}}\ \text{ACS}$$

In this scheme the set containing only a star {*} denotes a set of propositions with unknown truth value. A semantic abduction can encompass different kinds of inductive inferences. Affirming the consequent of a material implication is just one special case.

Definition 2 *Material abduction.* A material abduction is an inference that affirms the consequent of a material implication (AC), as a special case of a semantic abduction.

$$\frac{\text{Proposition P: } \{C\} \quad \text{Background B: } \{A \rightarrow C\} \cup \text{Hypothesis H: } \{A\} \models \text{P: } \{C\}}{\text{Hypothesis H: } \{A^*\}} \; \text{ACS: } \{AC\}$$

The material implication A → C could either be part of the hypothesis or belong to the established background assumptions B which should then be part of the antecedent of the semantic implication. Affirming the consequent of a material implication (AC) is the typical example of a semantic abduction. But the other discussed inductive inferences can be an instance as well, *i.e.*: denial of the antecedent (DA); division, attributing properties of wholes to parts (CMP); inferring causality between co-occurring events (CAU); and generalization from particulars to groups (GEN); see Table 4.3.

Explanation (ACS)	Proposition P	Background B	Hypothesis H
AC	C(y)	A(x) → C(y)	A(x) *
DA	Not A(x)	A(x) → C(y)	Not C(y) *
DVS	A(p) → C(p)	p part of w	A(p) → C(w) *
CAU	C(y)	A(x)	A(x) → C(y) *
GEN	$A(i_2) \rightarrow C(i_2)$	$A(i_1) \rightarrow C(i_1)$	∀x A(x) → C(x) *

Table 4.3: Some examples of explanation as semantic abduction (ACS): given $B \cup H \models P$, proposition P affirms the consequent to infer H.

But if inferences with material and semantic implications are part and parcel of abductive reasoning then we do not have an reason why the hypotheses of Calvin's father and Galileo differ. When is an abductive inference a scientific explanation?

4.5 Formation

There are in fact two very distinct ways to understand the terms "abductive inference" and "scientific explanation". In the first way the term is a verb and in the second way it is a noun. In the former sense it refers to the process of inferring and explaining. In the latter sense it refers to the product of that process. Abductive inference as defined by Peirce is first of all a process of inference. You assume two premises, and the conclusion of the inference is an explanation that could be correct.

But how do you know what specific hypothesis to infer? You could logically infer many different possible hypotheses that all would imply a surprising observation. (Why does the furnace not work? Is the switch broken? Is the gas pipe fractured? Oh wait a minute, did I pay my bill?) And on the other hand, coming up with only a single explanation that would non-trivially imply all our observations is no trivial exercise. Peirce's abductive inference scheme tells us nothing about what specific hypothesis to infer. He said: "The abductive suggestion comes to us as a flash" (1958, 5.181). His scheme only tells us under what condition to infer a statement as a hypothesis.

In the 1930's the philosopher Hans Reichenbach (1938) suggested that logicians should only address the problem of the nature of scientific theories and of their evaluation. The search and formation of new theories was taken to be an erratic and non-rational process that was not open nor relevant for a logical inquiry of knowledge. He suggested a distinction between a context of discovery and a context of justification in the study of scientific knowledge. This served as a demarcation of the problems relevant for epistemology. The study of the formation and discovery of hypotheses should be a problem for psychology. So according to Reichenbach's claim, logic should be able to evaluate a scientific explanation, regardless of how a hypothesis was inductively inferred or conceived. A good scientific explanation should satisfy certain logical conditions. One of those we already encountered: an explanation should logically imply the surprising observation. By its definition we already are sure that an abductive conclusion satisfies that condition. But both the explanation given by Galileo and that given by Calvin's father do so. So the question remains: what other conditions make an explanation scientific?

4.6 Explanation

Philosophers of science have long thought about the nature of a good scientific hypothesis. They set up certain conditions that would mark a valid and potentially successful explanation. We saw that any proper explanation should deduce a proposition from the explaining assumption. This is, by definition, possible if H combined with background assumptions B semantically implies P.

The set P may contain particular propositions, such as the proposition that certain objects have certain properties at a certain time. It can also contain general propositions, such as the proposition that all objects of a certain kind have a certain property, or that some object will have a certain property at a certain time. The background set B and hypothesis set H may also contain both particular and general propositions. General propositions in H and B can imply another general proposition in P. Together with an assumption about a particular they can imply particular propositions in P. In empirical sciences explanations are sought for particular or general facts about the world that are observed or assumed to be true. We will use the set O to refer to propositions about the world that are regarded to be certain because they are observed, given some criterion of proper observation.

In philosophy of science several conditions for a proper scientific explanation are proposed (*cf.* Aliseda-LLera, 1997, Flach, 1995). We will introduce some of them. There are both conditions for the explaining hypothesis and for the explained proposition. Given background assumptions B, proposition P, observations O; hypothesis H properly explains P if:

Conditions for the explaining hypothesis H:

HC_1. Implication:	$B \cup H \models P$
HC_2. Consistency:	H is compatible with B
HC_3. Non-triviality:	$H \not\models P$
HC_4. Simplicity:	H is minimal among the H's were $B \cup H' \models P$

Conditions for a proposition P that needs to be explained:

PC_1. Observation:	P is assumed to be true
PC_2. Novelty:	$B \not\models P$
PC_3. Anomaly:	$B \models \text{not } P$
PC_4. Indifference:	$B \not\models P$ and $B \not\models \text{not } P$

If PC_1 and any of the conditions PC_2 to PC_4 hold for a proposition, a hypothesis is required for which all conditions HC_1 to HC_4 hold. These are considered to be ideal conditions, proposed and defended by different logicians. Let us go through them and at the same time see whether the heliocentric hypothesis of Galileo and the hot air hypothesis of Calvin's father satisfy them:

Heliocentric hypothesis:	H: {center(sun)} $\models$ {phases(Venus)}
Hot air hypothesis:	H: {air cools $\rightarrow$ sun sets}

We already encountered the first condition HC_1. It dictates that an explanation of P consists in a deductive inference of P from B and hypothesis H. The philosopher Carl Hempel (1965) calls this hypothetical-deductive inference. By this condition an explanation consists of either a denial of the consequent of a hypothesis (DCH) or an affirmation of the antecedent of a hypothesis (AAH):

Background B: {A}		Background B: {not C}	
Hypothesis H: {A $\rightarrow$ C}		Hypothesis H: {A $\rightarrow$ C}	
———————————	AAH	———————————	DCH
Proposition P: {C}		Proposition P: {not A}	

In this way if B and H are true then they explain P. If P is true then it confirms H assuming B. Both the heliocentric and the hot air hypotheses comply as we saw earlier in our discussion in Section 4.4.

The second condition (HC_2) dictates that implications of $B \cup H$ should not contradict each other. That means that in case of contradiction either H or B should be substituted by a different set of propositions. The implication of the hot air hypothesis appears consistent with our other assumptions. However, the Heliocentric hypothesis contradicts the assumptions of Ptolemy, which were part of the background knowledge that, in Galileo's time, was assumed to be true. Condition HC_3 is meant to prevent the use of ad hoc hypotheses. It dictates that an observed proposition should not solely follow from the hypothesis. It should at least depend on some other assumptions that are not purely hypothetical. Both hypotheses comply. The fourth condition makes some requirements about the complexity of the hypothesis, given some interpretation of “minimal”. Both hypotheses do not seem unnecessarily complex.

In the next part of this thesis when we look at scientific practice, we will see that usually no employed hypothesis complies with all four conditions. It is usually argued that this fact does not mean that those hypotheses are unscientific or that the conditions are wrong. It is rather argued that the conditions define an ideal to be approached by science, given some justification for the conditions.

Now let us turn to the conditions for the explained proposition. Condition PC_1 states the assumption that a hypothesis in empirical science explains observations. If a consequence of a hypothesis is not observed, or on some other grounds certain to be true, then there is nothing to explain. While the four conditions for a hypothesis are each of them desirable, conditions PC_2, PC_3 and PC_4 are disjunctive; only one needs to apply. PC_2 states that a proposition only needs an explanation by a hypothesis H if it is not implied by what we already assume. PC_3 states that the observed proposition is in contradiction with the implications of our earlier assumptions. Or the background could be totally indifferent about it, as stated by PC_4.

The phases of Venus were a real anomaly (PC_3) for the assumption of Ptolemy. So by these conditions it required an explanation, which was properly provided by the heliocentric hypothesis. Yet, together with the assumption that the earth evolves around its axis, the rising of the sun is already explained by that hypothesis. It did not need another explanation. But logically there are always more explanations possible. So again, what makes the former a better explanation than the latter?

4.7 Prediction

Karl Popper contended that an explanation is no scientific explanation if it cannot be tested. He maintained that, before anything else, scientific reasoning is the systematic search for errors in our assumptions. Peirce also argued that therefore a proper explanation should at least predict propositions that are either novel, anomalous, or indifferent with respect to current (theoretical) assumptions. It should predict a P that satisfies conditions PC_2, PC_3, or PC_4, but not PC_1. Many wrong hypotheses may explain given observations, but true hypotheses will always correctly predict a new unobserved fact.

Logically a prediction of a proposition can be considered to be the same as an explanation, it should deductively follow from the hypothesis and background assumptions. But just as in the case of the definition of abduction we can make a distinction between affirming the antecedent of a material implication (AAH) or of a semantic implication (AAS). The former can again be part of the latter:

Definition 3 *Semantic prediction.* A semantic prediction is an inference that affirms the antecedent of a semantic implication (AAS). Given the antecedent $B \cup H$ that semantically implies P, the affirmation of the antecedent infers prediction P. Affirming the antecedent of an hypothetical material implication (AAH) is the prototypical example:

Background B: {A}
Hypothesis H: {A → C}
B: {A} ∪ H: {A → C} ⊨ P: {C}
———————————————————— AAS: {AAH}
Proposition P: {C}

We can consider the affirmation of the antecedent of a semantic implication as the general definition of prediction. Affirming the antecedent of a hypothetical material implication (AAH) is the prototypical AAS that provides the best bait for catching the truth value of an hypotheses by testing its prediction in the pond of nature. It is the ace of hypothesis testing. But others can be possible as well. A complete typology would be:

AAH: affirming the antecedent of a hypothesis
DCH: denying the consequent of a hypothesis
DAH: denying the antecedent of a hypothesis
ACH: affirming the consequent of a hypothesis

HAA: hypothetically affirming the antecedent of a background assumption
HDC: hypothetically denying the consequent of a background assumption
HDA: hypothetically denying the antecedent of a background assumption
HAC: hypothetically affirming the consequent of a background assumption

The value of a prediction for a hypothesis can be measured by the information we gain if we find out that the prediction comes true. We can call this its strength. In case of AAH a background assumption affirms the antecedent of a hypothetical implication. One infers the strongest prediction, its truth value either confirms or refutes a hypothesis. It is also possible to hypothetically affirm the antecedent of a hypothesis in the background assumptions (HAA). This is weaker because if the prediction P is observed it will not inform you about the truth of the hypothesis. But if not P is true it will refute the hypothesis, see Table 4.4 for all types.

Prediction (AAS)	Background B	Hypothesis H	Prediction P	If P is true then H is?	If P is false then H is?
AAH	A	A → C	C	Confirmed	Refuted
ACH	C	A → C	A *	Confirmed	Confirmed
DCH	Not C	A → C	Not A	Confirmed !	Refuted
DAH	Not A	A → C	Not C *	Confirmed !	Confirmed
HAA	A → C	A	C	?	Refuted
HAC	A → C	C	A *	Confirmed	?
HDC	A → C	Not C	Not A	?	Refuted
HDA	A → C	Not A	Not C *	Confirmed	?

Table 4.4: Types of prediction (AAS) of different strength: Given B ∪ H ⊨ P, background B affirms the antecedent of hypothesis H to infer prediction P.

So the route from theory to experiment is determined logically by an informative prediction that can be tested. The strongest test, the one that provides the most information, is always preferable. But there can be pragmatic problems to test it. The first problem is whether it is possible to observe the predicted property of a phenomenon. If not, the prediction is useless as an empirical test for the hypothesis. Most effort in the defense of the heliocentric hypothesis for Galileo was put in constructing a strong enough telescope to observe the predicted phases of Venus.

Some predictions state a possibility that will not naturally occur. But can you create an intervention such that the initial conditions for the possibility are forced? This is not always possible. The latest technology often makes observations and interventions possible that lay beyond our reach or sight without it. This makes technology an epistemological factor. Other predictions can easily be observed but will never occur according to the hypothesis. If they do not, how will you know they never will? Here lies the main problem of the hot air hypothesis.

P_9	If the air is hot the sun rises *	Hot(air) $\Rightarrow$ Rises(sun) *

This hypothesis logically implies that either:

P_{16}	The air cools and the sun sets	Cools(air) & Sets(sun)
P_{17}	The air is hot and the sun rises	Hot(air) & Rises(sun)
P_{18}	The air cools and the sun rises	Cools(air) & Rises(sun)

This is consistent with all our observations. But it also implies that it will never be so that the antecedent is true and the consequent is false, *i.e.*:

P_{19}	The air is hot and the sun sets	Hot(air) & Sets(sun)

This is its only test opportunity, that is unobserved so far. So, the only way to test the hypothesis is to create a situation where the air is kept hot by an intervention, and wait for the sun not to set. But how can we do that? The hypothesis is testable in theory, but not in practice. But does that make it an unscientific hypothesis?

4.8 Comparison

According to Theo Kuipers (Kuipers 2000) the question about the rationality of scientific reasoning is not only what it means to have a good scientific explanation, but also what it takes to have a better one. In this approach it is evaluated how one explanation compares to another. The best hypothesis would imply all true propositions about a domain. But acknowledging that this is the ideal goal, the value of an hypothesis is measured by how far it might be away from that goal in comparison with another hypothesis. A hypothesis that includes more true propositions then a competitor and has less counterexamples might be closer to the truth. This intuition is formalized in a rule of success. This inference rule is not deductive in nature, but abductive. If the more successful theory would be closer to the truth that would explain why it is more successful. In this light Calvin's father's explanation is not so much unscientific, but just not as good as Galileo's, because next to explaining the phases of Venus, it also explains other phenomena such as stellar parallax. Yet there are more conditions formulated that characterize a good scientific explanation. In Chapter 6 I discuss how one of them, the simplicity of a theory, is related to the probability of its predictions.

4.9 Conclusion

In this chapter I asked the general question: what is the rational use of theory and experiment in the process of scientific discovery, as proposed in the study of logic? More specifically I looked at logical prescriptions for scientific theories and scientific reasoning. To address these topics I discussed an illustrated example that contains a series of inferences that are marked as fallacies from the viewpoint of logic and argumentation. Yet I argued that these inferences are common in science and part of abductive inference as defined by C.S. Peirce. I further made a category distinction between semantic abduction and material abduction. I argued that the latter, as well as other types of inductive inference, constitute a special type of the former under this definition.

I first discussed the validity of deduction, induction, and more specifically abduction in scientific reasoning. Scientific reasoning includes inferences about hypotheses of which we do not or cannot know whether they are true. What logic tells us most importantly is what a valid inference looks like. It defines under what conditions we can safely accept the conclusion of an argument. In the case of deduction we know that the conclusion is true when the premises are true. In the case of abduction or explanation we can know that the premises are true, but we have no guarantee for the conclusion. What valid reasoning can do is check whether the conclusion of an inference satisfies certain conditions. For explanation it can check whether a hypothesis is *e.g.* successful, non-trivial or consistent. But these are ideal conditions that still do not determine its truth. Yet they may be functional for establishing its similarity to the truth. I argued that prediction is not just deduction. A good prediction with the aim to test a hypothesis should satisfy other conditions as well.

In sum, what is rationality in scientific discovery? According to logic scientific discovery is a process of observing, describing, explaining, predicting and intervening in natural phenomena. A phenomenon is empirically discovered by observing it in the world. An explanation of that phenomenon may predict the existence of other phenomena that could be observed or created by a specific intervention in an experiment to test that prediction. As an answer the specific questions of this thesis from Section 1.3, we may not that according to studies in logic the following holds:

Question 1 What is the structure of a scientific theory? Theories are logically represented as a set of hypothetical propositions H that together with propositions describing background assumptions B semantically imply the propositional facts P they explain, *i.e.* $B \cup H \models P$.

Question 2 What is the process of scientific reasoning? The process of reasoning is different for the explanation and prediction of facts, see Table 4.5.

Problem	Premise	Background	Inference	Conclusion	Properties
Explanation	P	B	Abduction	H: {*}	$B \cup H: \{*\} \models P$ H is minimal
Prediction	H	B	Deduction	P: {*}	$B \cup H \models P: \{*\}$ P is informative

Table 4.5: Short overview of the inference types discussed in this chapter

Explanation of a phenomenon involves the abduction of a simple hypothesis from which the properties of an observed instance of that phenomenon can de deduced. Induction, as conceived as the generalization from the property of one instance of a category to all instances, is in this sense a special kind of abduction. Prediction involves the deduction of informative consequences from a given hypothesis.

Question 3 What is the route between theory and experiment? The route between theory and experiment typically involves six steps (explanation follows):

1. Observation of a phenomenon P: observe p_m and p_n
2. Description of P: $P: \{A(p_m) \rightarrow C(p_n)\}$
3. Explanation of p by a new hypothesis: $B \cup H: \{*\} \models P$
4. Prediction by a hypothesis: $B \cup H \models P: \{A(p_i) \rightarrow C(p_j)\ *\}$
5. Intervention in an experiment: create $A(p_i)$
6. Observation in an experiment: observe p_j

An observation of a phenomenon p in step 1. consists in observing natural objects such as *e.g.* p_m and p_n. The description of p in step 2. consist in categorizing the properties of the phenomenon, *e.g.* in A and C, and making a statement about those properties, *e.g.* $A \rightarrow C$. After finding an explanation, in step 3., that implies that statement, a prediction could be deduced in step 4. This prediction can include that if an object p_i has property A, then object p_j will have property C. In step 5. the situation $A(p_i)$ can be forced by an intervening experiment. The last step, observing the consequence of the intervention, closes the circle by being of the same kind as the first step. The experimental discovery of the truth value of the prediction either refutes or confirms the hypothesis (or a background assumption). A more advanced logical approach can evaluate an hypothesis by comparing its success with that of competing hypotheses.

In the next chapter I will discuss rationality in the process of scientific discovery in terms of the study of cognition. In this approach rationality can be understood as part of learning to solve problems heuristically.

* * * *