The Referential Basis of Abduction and Logic

Erich Prem[1]

Abstract. AI has often mistreated logic and abduction. A closer view at the historical roots of logic and at the system theoretic structure of logical modeling reveals that the foundations of logic lie in a science of argumentation rather than reasoning. The basis of such a science is a clarifcation of the term "argument" that also leads to a reduction of the semantics of natural language. In logic, the reductions are tacit semantic restrictions to specialized meanings of terms and forms of conceptual relations as they are used in arguments. These relations are the conditions of the possibility of conceptual reference and at the same time of logical reasoning, where abduction and the identification of something as something form the basis of both.

1 Introduction

The continuing argument concerning the nature of abductive reasoning in Artificial Intelligence (AI) is symptomatic for the poor understanding of the philosophical basis of logic that prevails in AI. Many, though not all, AI researchers have insisted on or failed to question the claim that logic is a "science of reasoning". A fundamental misunderstanding concerns the term "reason", which connotes thought and thinking as well as ground and cause. Consequently, logical theories and the computational efforts of implementing them as parts of AI systems have often been mistreated as theories concerning the nature of human thinking, the thinkable or, even worse, as ontological conceptions.

Logical laws have therefore held an unclear status somewhere between mathematical axioms, philosophical trivialities, psychological regularities, or ontologically immanent truths. Only rarely (as in [12]) have they been recognized as what they really are: tacit conditions for the applicability of different logical calculi that work as semantic reductions of linguistic terms and strucures.

2 Logic and AI

In a conventional AI program the sole purpose of the knowledge base and the inferential procedure working on it is the deduction (or rather calculation) of truths about some natural system which is to be modeled. Reduced versions of predicate logic are used to model "knowledge about the world" (see fig. 1). Entities in the natural world are mapped on formal symbols in F, mostly through the use of measurement devices. Formal rules operating in F lead to the derivation of other symbols that can be interpreted (again using meters) as being about N.

Of course, a precise account of the way the world is modeled would have to distinguish between the world, knowledge about it, and its formal description. Fig. 2 shows this distinction and includes a further step: the logical modeling of aspects of language.

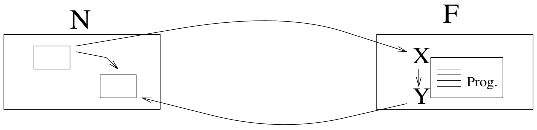

Figure 1. The deduction of true sentences about a natural system N in a formal system F.

Every step to the right in fig. 2 denotes a process of abstraction (arrows number 1, 2, and 3), which encods elements in symbols. Conversely, every arrow pointing to the left means a decoding process. Arrow number 4 depicts the process of logical inference which typically results in the computation of a statement about the world (6). (Arrows 7, 8, and 5 only have an indirect interpretation, e.g. 5 means that derivations of logical programs must stand the test in reality.)

3 Logic and Argument

In this picture logic operates on formal aspects (symbolically encoded elements) of language, not of knowledge or the world. Therefore, in this picture, logic is a science of argumentation, not of reasoning.[2] This is exactly how it has been developed by Plato: as a tool against mere sophistry. His discussion of what a real argument is, has been refined later by Aristotle towards a "demonstrative science".

> We must first state the subject of our inquiry and the faculty to which it belongs: its subject is demonstration and the faculty that carries it out demonstrative science. [1, 24]

Aristotle's introductory statements do not bear the character of explaining laws of language (or thought!), but of scientific reductions of the semantics used in the context of "correct" arguments. In the following example, Aristotle defines the form of a premise. This is a definition, *not* an explanation.

> A premise then is a sentence affirming or denying one thing of another. This is either universal or particular or indefinite. [1, 24]

The restrictions on the meaning of terms are part of a methodology that helps in avoiding misunderstandings. For example, the "principle of contradiction" is part of such a methodology by stipulating that statements about the world should not be true and false at the same time. Other axioms of this methodology ensure that statements are only true or false, etc. A modern example of such an implicit reduction (compared to the meaning of sentences and words in spoken natural language) used in the context of AI is the following:

[1] The Austrian Research Institute for Artificial Intelligence, Schottengasse 3, A-1010 Wien, Austria. (erich@ai.univie.ac.at)

[2] *Argumentation* is used in the sense of proofing, explaining, making the case, not in the sense of persuading or talking round.

ECAI 96, *12th European Conference on Artificial Intelligence*
Workshop on Abductive and Inductive Reasoning
Budapest, Hungary.

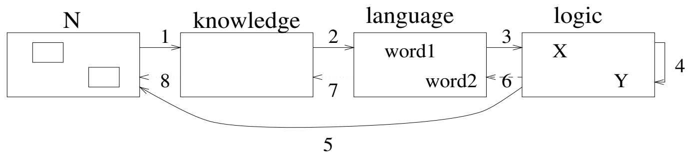

Figure 2. Phases of modeling in knowledge-based systems.

> Even if we do not know whether there is life on the planet Jupiter, we know that the sentence
>
> > "There are monkeys on Jupiter or there are no monkeys on Jupiter."
>
> is true. The truth of the sentence can be determined from its structure alone, without knowing whether its consituents are true or false. [5]

The reduction implicit in this example consists in stripping away any common sense connotations or usages of such sentences with the aim of getting a formalizable instance of the abstract rule *"P or not P."* Manna & Waldinger's claim is, again, not an explanation, but an implicit definition of semantical rules about the meaning of terms such as "or", "is", etc. The example adheres to the "principle of contradiction," which is not a statement about what can be thought or about possible state of affairs in the world. It is a requirement for a "reasonable" argument (and at the same time the implicit definition of what it means for an argument to be reasonable).

The step to such a linguistic view of logic is necessary to give a coherent account of ab-, in- and deduction. In what follows, I will show that the semantic restrictions in logic are (designed as) the conditions of the possibility of any conceptual (i.e. notional) reference. The very idea of a science of argumentation is based on a mathematical understanding of the relation between sign and referent. In Plato's case, a sentence like "Excellency is knowledge." is used for the clarification of relations among terms. Plato uses the metaphor of canvas to clarify these relations which results in a simple (but clear) view of how symbols relate to the world and to each other. (His canvas have been replaced by Venn diagrams today [10].)

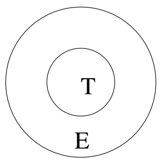

Figure 3. Platon's use of canvas to clarify the relations of concepts: Excellency (T) is knowledge (E).

Another technique to avoid misunderstandings and invalid arguments is based on hierarchical concept trees ("arbores porphyrianae"). This is the method that has been improved by Aristotle, e.g. as a means for the natural scientist (biologist) to structure the world of living beings. The true advantage of his method, however, lies in the fact that the neat structure of dichotomies allows for a study of the relation of terms in such a tree. *If* concepts can be arranged in a hierarchical (e.g. binary) tree *then (and only then!)* the general relations among concepts and to sentences (All A are B.) can be investigated systematically. Therefore logic, in its original form is a tool to avoid terminological confusion and invalid arguments. It achieves this at the cost of reducing the semantics of words and sentences. But as an additional feature, we can now study conceptual relations and the validity of arguments which use these relations only with respect to their form while disregarding the content completely. However, the existence of such a dichotomic tree is the necessary prerequisite for making such a study possible. The tree remains as the background structure which must exist before any logical inference can be made.

In order to find out more about logical reasoning (and in- and abduction), it is necessary to clarify what the aforementioned semantic restrictions exactly are and how they influence the structure of logical arguments. I continue by arguing that the conditions of the semantic reductions have been identified with the possibility of any conceptual reference in the area of transcendental logic.

4 Conditions of Conceptual Reference

According to work in transcendental logic [13] the necessary prerequsites for conceptual reference are

1. existence (principle of identity, A is A),
2. discrimination (A is not non-A),
3. intendability
   (a) by reference to something (A is B)
   (b) that is distinguished (if B then non-B)
   (c) exlusively (A is either B or non B).

These principles[3] arise naturally from the view depicted in fig. 1. Steps 1 and 2 are necessary to cut out entities in the natural world, while steps (a) through (c) explain the character of this mapping and the formal aspects more clearly. Some of these principles are known to logicians as principles of logic.

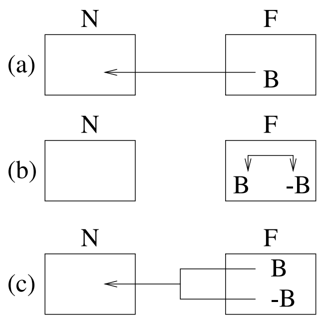

Figure 4. Conditions (a)-(c) of conceptual reference, i.e. formal reference to "natural" objects.

[3] The philosophically trained reader will realize the similarity to Kant's categories [4].

Fig. 4 depicts how principles (a) through (c) could be imagined intuitively. It can be seen, how the principles ensure a clear meaning of the linguistic elements in F. The signs of F refer to one element in N. For every such element (B) in F there is also its counterpart (not-B). And, finally, every element in N refers either to B or not-B, *tertium non datur.*

Not only are these conditions the necessary prerequisites for reference in general, they are also the description of terminological relations in concept trees, simply because symbols can also refer to other symbols. This application of the principles of conceptual reference to concepts themselves naturally leads to the construction of hierarchical and dichotomic concept trees, i.e. to the background structure already assumed before logical propositions can be investigated. If we want to construct a science of argumentation, which due to its very nature, must operate on linguistic structures, we need to ensure a similar clarity as in the case where signs refer to natural objects for the relation among words. Fig. 5 suggests how such an application leads to the construction of concept trees.

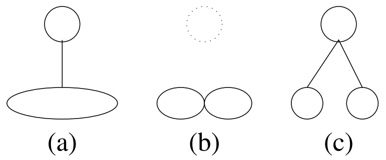

Figure 5. Conditions (a)–(c) of conceptual reference applied to formal concepts.

Following from our considerations of conceptual reference, the dichotomic concept tree represents the kind of background structure necessary for making clear statements about the world and, more importantly, about statements (since connected statements becom arguments). A science of argumentation leads to a specific, mathematically oriented view of language (as mapping) and (linguistic) reference as mapping of a certain kind. *If* concepts and their corresponding states of affairs) are of a specific structure, *then* I can use the form of the relations among concepts to study the laws of "correct argumentation". "Correct" here means correct with respect to the state of affairs. These considerations lead to a basically deductive formal structure, the concept tree, as the background assumption of all logical reasoning.

The close relation of argumentation to logical reasoning has, of course, been realized by logicians and researchers in the field of AI also. [2, 7, 8] and others have made a step in this direction by arguing that reasoning *with incomplete information* rests on some "principle of argumentation." However, to my knowledge, no in-depth clarification of how these reasons and the act of "giving reasons" relate to language or symbols in logical models exists. It remains doubtful whether there can be a "theory of rational inference" (as Pollock suggests) without such a clarification. None of the above authors realizes neither that logic *is* (a science of) argumentation nor that logic implicitly *defines* what a correct argument is.

The reason that this "transcendental" property of logic remains concealed lies in the fact that logic is mostly used to infer facts from other facts, so as to generate *new* knowledge and to study rules that make such inference possible. However, and this is the central claim of this paper, the development and study of such rules already rests on the strict background structure of a deductive system of concepts. In order to support this claim, we shall now study the logical inference procedures and their relation to this background structure.

5 Abduction and Induction

5.1 Abduction

Abduction is usually understood as the process of generating explanations or hypotheses for observed facts, often with respect to a background theory. Induction can be defined as consisting in

> identifying the similarities between observed objects or events, and hypothetically extending those similarities to unobserved objects or future events. [3, p.1]

The rather vague difference between the "generation of a hypothesis" and the "hypothetical extension of similarities" has been the source of many confusions concerning the true nature of in- and abduction. This is all the more true since both forms of reasoning are considered to generate "new knowledge" from observations, unlike deduction.

A prototypical example of an abductive inference has been given by Charles S. Peirce.

All beans in this sack are white.
These beans are white.
$\Rightarrow$ These beans are from this sack.

This argument is based on a rule (or "background theory") *"All beans in this sack are white."* and a result or "surprising observation" [6, 3:325f], namely that *"These beans are white"*. The conclusion that *"These beans are from this sack."* always happens without a guarantee, but this is necessary to discover new knowledge. Other names for this form of reasoning are "retroduction" or simply "hypothesis" (Peirce). It is clear that nothing forces us to draw the above conclusion. The beans could equally well stem from another sack than the one which holds only white beans. Let us therefore compare Peirce's example with another example given by Aristotle.

All stars glow by their own light.
This object does not glow by its own light.
$\Rightarrow$ This object is not a star.

The form of both arguments nearly is the same. The main difference is the negated second premise and the negated consequence. The position of the terminus medius in this process suggests that such an inference corresponds to Aristotle's second figure [1, 26b/27a]. The terminus medius is at the position of the predicates, it is what is "said" about both subjects. In a *formal* approach that is only interested in drawing valid conclusions, abduction is only possible with a negative conclusion.

5.2 Induction

The typical example for induction consists in enumerating observations and generalizing the observation to a rule, for example:

My dog has four legs.
Your dog has four legs.
This dog over there has four legs.
$\Rightarrow$ All dogs have four legs.

Induction is used to inform us about the total set of objects with a specific property. Generally, in induction we are given a set X with elements x_i and a subset S of X. We know now (e.g. from experiments with some s_i) that for all s in S there holds a property $P(S)$. Induction means to generalize from $(\forall s \in S)P(s)$ and $S \subset X$ to $(\forall x \in X)P(x)$. A simpler example for such an argument is the following one:

Philosophers are mortal. $(\forall s \in S)P(s)$
Philosophers are humans. $S \subset X$
$\Rightarrow$ Humans are mortal. $(\forall x \in X)P(x)$

This example also has a counterpart in the form of an Aristotelian syllogism (taken from [10]).

Some humans are not greeks.
All humans are mortal.
$\Rightarrow$ Some mortals are not greeks.

Again, the form of this argument is essentially the same as above, but with only a particluar conclusion and a negated first premise.

5.3 Negotiation

According to Aristotle, all questions are questions about the middle term in a syllogism. Let us therefore ask, what the features in the middle terms are in in- and abduction.

In the case of abduction, it is the posession of a common property that negotiates between the premises and the conclusion. In abduction there is one concept at the position of the predicate that mediates between the two subjects: B is C. A is C. $\rightarrow$ A is B.[4] It is the predicate that allows for such an analogy.

Drawing the argument in the form of a tree yields fig. 6, which represents a part of a hierarchical concept tree.

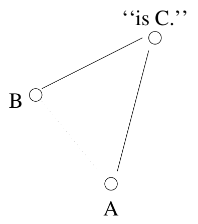

Figure 6. Part of arbor porphyriana. The predicate "is C" is drawn at the position of the general. Straight lines show relations given, dotted line represents conclusion.

This example allows to explain why I have termed the hierarchical concept tree to be the "background" structure. The left part of fig. 7 shows fig. 6 as a part of a larger dichotomy.

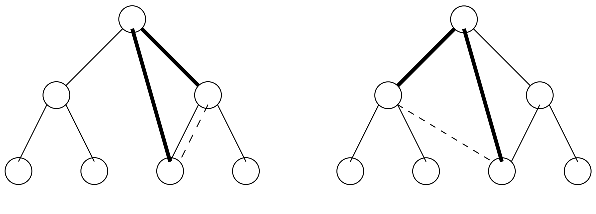

Figure 7. Left: Valid abduction. Right: Invalid abduction. Thick lines represent premises, dotted lines are conclusions.

The inference from "All B are C." and "All A are C." to "All A are B." is valid only in the left, but not in the right tree. This is why abduction is valid only with a negative conclusion. Despite the fact that both premises are true *and* the background structure is a dichotomic tree, we might actually conclude something wrong.

[4] More precisely wrt to the example by Peirce: All B are C. A is C. $\rightarrow$ A is B.

Similarily, in Aristotelian induction the terminus maior is proven for the medius by means of the terminus minor [1, 68b]. In Aristotle's third figure the terminus medius is the subject of both predicates: A is B, A is C; i.e. in the hierarchical concept tree, we find the medius at the bottom (cf. right tree in fig. 9).

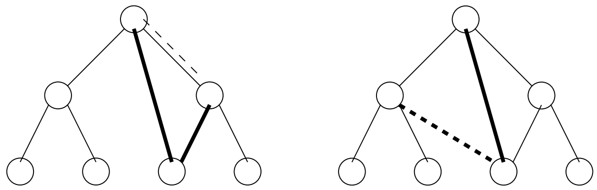

Figure 8. Left: Valid abduction. Right: No completely invalid conclusions are possible, at least particular conclusions can be drawn. Thick lines represent premises, thin dotted line is a valid conclusion.

Fig. 8 explains what the background structure of the dichtomic tree means for induction. The left tree depicts a valid form of argumentation. The right tree suggests that there is no way to have the same problem as in abduction. Two valid inductive premises will always allow to draw at least a particular conclusion, because the thick dotted line in fig. 8 cannot represent a valid premise, if the thick line represents one, since opposite nodes at the same level correspond to negated concepts.[5]

Fig. 9 depicts the relations of ab-, in- and deduction in terms of a part of the binary concept tree and in terms of inclusion relations among sets.

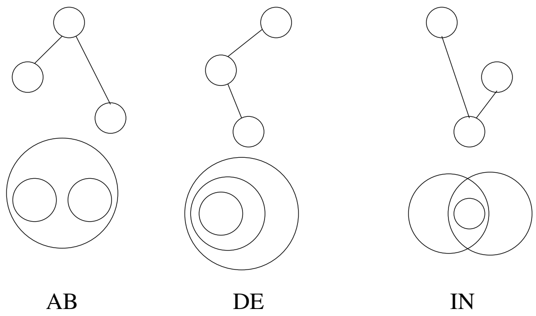

Figure 9. The relation of terms in the arbor porphyriana in the case of ab-, de-, and induction. Corresponding inclusion relations among sets are depicted below the trees.

It is worthwhile pointing out the different kind of inclusion relations and the corresponding principles of conceptual reference. Abduction, as drawn in fig. 9 (right), depicts the relation of inclusion, or *A is B.* Induction depicts exclusion, because one subject appears in two different (excluded) views: as *B and (another) not-B.* Finally, the perfect inference possible in deductive reasoning stems from the double role of the terminus medius as subject and predicate, which

[5] If this point remains unclear, the reader is invited to study all valid forms of Aristotle's third figure with respect to the tree, which are IAI, AII, OAO, and EIO. Cf. [10]

ensures perfect reference. This last two observations become clearer on examination of the referential properties of signs.

5.4 Signs

I have argued at length elsewhere [9] that the three cases of terminological relations in the concept tree are really in correspondance with the three figures in Aristotelian reasoning, with Kant's categories, and Peirce's three types of signs.

The abductive process of reasoning becomes represented in one special type of sign, the *icon*. It is the possession of a common (essential) property P which negotiates between the sign and referent. By reference to P, the icon represents *intension.* In more formal terms, let M be the set of object states, Q be the set of the sign states. Then it is the fact that both, M and Q, possess (belong to a set) P which makes them stand for each other. (M is P. Q is P. $\rightarrow$ M is Q.) This negotiation takes exactly the form of abduction or analogy.

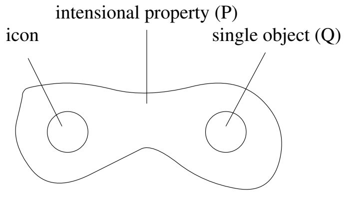

Figure 10. Negotiation through an essential property.

In order to accept our view that the icon represents abduction, it is necessary to see that the role of intension corresponds to the role of the general in classical reasoning; that the object comprises the singular and that the sign represents the special. Terminus medius of iconic representation is the icon's intension.

Identifying sign and object empirically will therefore succeed if both are recognized as being essentially of some property P. However, this reference will always happen without a guarantee. Neither can I be sure that the icon "really" means what it means to me, nor can I be sure that a word means what it means to me.

The *index* is a bit more problematic, mainly because we are not so much used to working with it. We remember Peirce' hint that the index is also index without anyone interpreting it. Therefore, what is characteristic about it, must lie in the object itself. This is what is originally meant by the "naturalness" of the index or, in other terms, that the connection exists because of natural law. It is the object alone which informs about its quality of being an index.

In the case of the index the single objects negotiate their meaning, they are the elements of induction, which are measured by a meter (the index, e.g. a thermometer) and which are instances of a general property, like temperature. Figure 11 depicts this relation (middle).

Finally, the symbol plays the role of subject and predicate. Its perfect or "ideal" reference is based on the assumption that a sign either refers to its object, or it does not, tertium non datur.

6 Discussion

After so many views and pictures (literally) of logic, it is probably time to reconsider what the aim of this paper is. The basic motivation for this in-depth inquiry into the nature of logic has been the feeling that the notions of ab- an induction lack clarity and preciseness. To

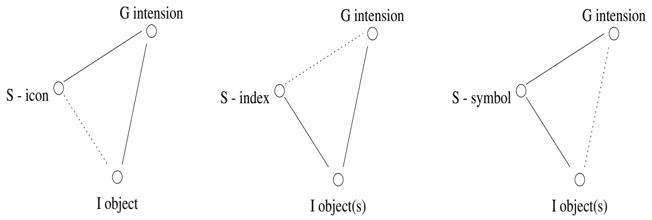

Figure 11. Parts of arbor porphyriana. G-general, S-special, I-singular. Straight lines show relations given, dotted line represents conclusion. Abduction (left), induction (middle), deduction (right).

my understanding, this deficiency is based on a wrong picture of what logic is.

If logic is considered to be a "science of truth" (Quine), the danger arises to regard the propositions of logic as ontological statements about what can or cannot be the case. If, on the other hand, logic is construed as a "science of reasoning" (Manna & Waldinger), then the danger arises that logical propositions are considered to concern the nature of human thought.

In the view outlined in this paper logic is constructed as a science of argumentation. In this framework, the form of an argument is based on the referential properties of signs in scientific languages, i.e. on the properties of the relation between formal elements used to describe natural or other formal entities. It remains true, of course, that such a turn in viewpoint can also lead to a study of "the laws of thought", if one believes that thinking must always already happen in terms of conceptual grasping of reality. In fact, this view has been widely adopted in the AI community and lead to the interpretation of thinking as "speaking to oneself" (to paraphrase Wittgenstein).

One could also be tempted to ask which form of reasoning is the basic one. With respect to the argument that logic is based on the clearcut properties of reference, it should be clear that this basic form is abduction. Abduction serves to connect A with B and is always already used when refering to something by means of words. Abduction is used in the "is" of any reference. "This is a dog!" is a statement which can only be made without a guarantee of being understood. It is exacly in this sense that abduction comprises a primordial form of reasoning.

> Deduction proves that something *must* be; Induction shows that something *actually* is operative; Abduction merely suggests that something *may be.*
> ([6, 5:171] quoted after [3].)

In this light it is not surprising that "in some sense" abduction is the basis for induction, as Peirce has already noted. Even a major proponent of deductive reasoning, Simon argues that abduction is the only place where the "real inventions" are made [11].

Abduction comprises the first and most fundamental form of categorical synthesis. Whereas it is generally held that "abduction is used to generate reason", I argue that abduction is reason. By representing the fundamental principle of terminological reference it forms the essence of logical inference due to to the linguistic nature of logic per se.

ACKNOWLEDGEMENTS

The Austrian Research Institute for Artificial Intelligence is sponsored by the Austrian Federal Ministry for Science, Transport, and the Arts. I would like to thank Johannes Fürnkranz, Gerhard Widmer and an anonymous reviewer for his helpful comments.

REFERENCES

1 Aristotle *Analytica Priora.*

2 Phan M. Dung, On the Acceptability of Arguments and its Fundamental Role in Nonmonotonic Reasoning and Logic Programming, in Bajcsy R.(ed.), Proceedings of the Thirteenth International Joint Conference on Artificial Intelligence, Morgan Kaufmann, San Mateo, CA, pp.852-859, (1993).

3 Peter A. Flach, *Conjectures: An Inquiry Concerning the Logic of Induction,* ITK Dissertation Series, Tilburg University, The Netherlands, 1995.

4 Immanuel Kant, *Kritik der reinen Vernunft,* (A critique of pure reason), 1781.

5 Z. Manna, R. Waldinger, *The Logical Basis for Computer Programming* Vol. 1 Deductive Reasoning, Addison-Wesley, Reading, MA, 1985.

6 M.H. Fisch, W. Kloesel (Eds.), *Writings of Charles Sanders Peirce. A Chronological Edition*, Indiana University Press, Bloomington, 1982.

7 John L. Pollock, OSCAR: A General Theory of Rationality, JETAI Journal of Experimental and Theoretical Artificial Intelligence, **1** (3), (1989).

8 John L. Pollock, How to Reason Defeasibly, Artificial Intelligence, **57** (1), (1992).

9 Erich Prem, Symbol Grounding and Transcendental Logic, in Niklasson L. & Boden M.(eds.), Current Trends in Connectionism, Lawrence Erlbaum, Hillsdale, NJ, p. 271–282, (1995).

10 Willard Van O. Quine, *Methods of Logic*, Rev. Ed., Holt, Rinehart and Wilson, NY, 1964.

11 Herbert A. Simon,‘ The Logic of Rational Decision’, In *The British Journal for the Philosophy of Science.* pp. 169–186, **63** (16), (1965).

12 Pirmin Stekeler-Weithofer, *Grundprobleme der Logik*, De Gruyter, Berlin, Grundlagen der Kommunikation, 1986.

13 Kurt W. Zeidler, *Grundriss der transzendentalen Logik*, Junghans Verlag Cuxhaven, 1992.