A PHILOSOPHICAL ANALYSIS OF COMPUTATIONAL MODELING IN COGNITIVE SCIENCE

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF INFORMATICS
OF
THE MIDDLE EAST TECHNICAL UNIVERSITY

BY

BURCU AYŞEN ÜRGEN

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE
IN
THE DEPARTMENT OF COGNITIVE SCIENCE

SEPTEMBER 2007

Approval of the Graduate School of Informatics

Prof. Dr. Nazife Baykal
Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science.

Prof. Dr. Deniz Zeyrek
Head of Department

This is to certify that we have read this thesis and that in our opinion it is fully adequate, in scope and quality, as a thesis for the degree of Master of Science.

Assoc. Prof. Dr. Cem Bozşahin Co-Supervisor	Assoc. Prof. Dr. Samet Bağçe Supervisor

Examining Committee Members

Assist. Prof. Dr. Bilge Say	(METU, COGS)	
Assoc. Prof. Dr. Samet Bağçe	(METU, PHIL)	
Assoc. Prof. Dr. Cem Bozşahin	(METU, CENG)	
Prof. Dr. Varol Akman	(Bilkent Univ., CS and PHIL)	
Assist. Prof. Dr. Joshua Cowley	(Bilkent Univ., PHIL)	

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this wok.

Name, Last Name: Burcu Ayşen Ürgen

Signature: ______________________

ABSTRACT

A PHILOSOPHICAL ANALYSIS OF COMPUTATIONAL MODELING IN COGNITIVE SCIENCE

Burcu Ayşen Ürgen
MS, Department of Cognitive Science
Supervisor: Assoc. Prof. Dr. Samet Bağçe
Co-supervisor: Assoc. Prof. Dr. Cem Bozşahin

September 2007, 43 pages

This study analyses the methodology of computational cognitive modeling as one of the ways of conducting research in cognitive science. The aim of the study is to provide an understanding of the place of computational cognitive models in understanding human cognition. Considering the vast number of computational cognitive models which have been just given to account for some cognitive phenomenon by solely simulating some experimental study and fitting to empirical data, a practice-oriented approach is adopted in this study to understand the work of the modeler, and accordingly to discover the potential of computational cognitive models, apart from their being simulation tools. In pursuit of this aim, a framework with a practice-oriented approach from the philosophy of science literature, which is Morgan & Morrison (1999)'s account, is employed on a case study. The framework emphasizes four key elements to understand the place of models in science, which are the construction of models, the function of models, the representation they provide, and the ways we learn from models. The case study

Soar (Simon, Newell & Klahr, 1991), is a model built with Soar cognitive architecture (Laird, Newell & Rosenbloom, 1987) which is representative of a class of computational cognitive models. Discussions are included for how to make generalizations for computational cognitive models out of this class, i.e. for models that are built with other modeling paradigms.

Keywords: Computational cognitive modeling, model, Soar, philosophy of science

ÖZ

BİLİŞSEL BİLİMLERDEKİ BİLGİSAYARLI MODELLEMENİN FELSEFİ BİR ANALİZİ

Burcu Ayşen Ürgen
Yüksek Lisans, Bilişsel Bilimler Bölümü
Tez Yöneticisi: Doç. Dr. Samet Bağçe
Ortak Tez Yöneticisi: Doç. Dr. Cem Bozşahin

Eylül 2007, 43 sayfa

Bu çalışma, bilişsel bilimlerdeki araştırma yöntemlerinden biri olan bilgisayarlı bilişsel modelleme metodolojisini incelemektedir. Çalışmanın amacı, insanin bilişini anlamada bilgisayarlı bilişsel modellerin nasıl bir yere sahip olduğunu gösteren bir anlayış sunmaktır. Yalnızca, belirli bir deneysel çalışmayı simule ederek ve deneysel verileri tekrarlayarak belirli bir bilişsel fenomen için görüş bildiren modellerin çokluğu düşünülerek, bu çalışmada modeli yaratan kişinin yaptığı işi anlamak ve böylece bilgisayarlı bilişsel modellerin simülasyon aracı olmalarının dışındaki potansiyellerinin keşfedilebilmek için pratiği anlamaya dayalı bir yaklaşım içinde bulunulmuştur. Bu amaç için bilim felsefesi literatüründen pratiği anlamaya dayalı bir yaklaşım olan Morgan ve Morrison (1999)'ın görüşü bir örnek model üzerinde uygulanmıştır. Bu görüş, modellerin bilimdeki yerini anlamak için dört temel eleman üzerinde vurgu yapmaktadır; modellerin kurulması, modellerin işleyişi, modellerin temsili ve modellerden öğrenme şekillerimiz. Üzerinde çalışılan örnek, Soar mimarisiyle yaratılmış bir model olan ve belirli bir grup modeli temsil eden Q-Soar modelidir. Çalışmada, bu grup modeller dışındaki bilgisayarlı bilişsel modeller için nasıl genellemeler yapılabileceği hakkında tartışmalar verilmiştir.

Anahtar kelimeler: Bilgisayarlı bilişsel modelleme, model, Soar, bilim felsefesi

To all I lost

ACKNOWLEDGMENTS

I express sincere appreciation to my principle thesis supervisor Assoc. Prof. Dr. Samet Bağçe for giving me latitude and encouragement throughout his guidance in this research. He has always been kind and insightful as an academics and as a person. I thank to my co-supervisor Assoc. Prof. Dr. Cem Bozşahin for his suggestions and comments. I am grateful to Assist. Prof. Dr. Bilge Say for her understanding in dealing with assistantship duties at difficult times and for her encouragement in many issues regarding academic life including this thesis work.

I am grateful to every member of my family who has dedicated himself/herself to provide me the best of everything. To my grandfather, Emin Ürgen, I specially thank, for providing me the unique inspiration with his way of living to stand firmly in life.

Finally, I am grateful to all the writers and poets who have their place in my small library for making life bearable.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

As a research field embracing various disciplines such as psychology, artificial intelligence, linguistics, philosophy and many other disciplines, different approaches are taken in conducting cognitive science research. One such approach is computational modeling of human cognitive phenomena**.** It is basically defined as development of computer models of cognitive processes, and the use of such models to simulate or predict human behavior (Cooper, 2002). Its history dates back to the information-processing revolution in 1950s and 1960s, which viewed thinking as manipulation of symbols and employed computer simulations to build theories of thinking (Newell & Simon, 1976). After its emergence computational modeling has gone under substantial development together with the introduction of different computational paradigms, e.g. connectionism and cognitive architectures in dealing with cognitive phenomena (Rumelhart & McClelland, 1986; Anderson & Lebiere, 1998; Newell, 1990; Kieras, 2004).

The main aim of this study is to discuss the role computational models play in cognitive science. Specifically, the question under investigation is, in what ways computational models contribute to understanding of human cognitive phenomena. Underlying my intended discussion on this topic is the belief that computational modeling is one of the major enterprises in cognitive science research for the aim of obtaining scientific knowledge we have regarding cognitive phenomena, and that one needs to understand how the procedures applied in scientific practices enable science to produce and present its findings as scientific knowledge. Considering the presence of computational cognitive models which have been just given as a simulation of an experimental study, replicating the empirical results, without giving any construction detail or possible further uses of the model, an evaluation of the practice is necessary to see what the place of this practice is among other ways of conducting cognitive science research (e.g. experimental studies).

In pursuit of this aim, I appealed to the literature of the discipline of philosophy of science since it takes scientific practice as an object of investigation and accordingly provides ways of examination of the scientific practices. Specifically, I appealed to its literature focusing on the concept of *model,* considering it to be enlightening for me to provide an understanding of computational cognitive modeling *as a practice* in pursuit of the research question under investigation. I came up with various accounts on models during the literature review within the philosophy of science, which may be classified basically into two groups; one that views model as a component of science but subsidiary to the primary component theory, and another that views model as a product of a special *practice*, modeling, which is thought to deserve a distinct treatment from theorizing, and provides a framework to understand that special practice. The former view is represented by two major sub views called Received View (Campbell, 1920; Carnap, 1939; Hempel, 1965; Nagel, 1979) and Semantic View (Suppe, 1977, 1989; van Fraasen, 1980, Giere, 1988) in the philosophy of science literature. The latter view is a more recent one and has been developed by Morgan and Morrison (1999) by studying many case studies of modeling from various natural and social sciences. Morgan and Morrison's (1999) view seems to be an appropriate tool to make benefit from for the aims of this thesis since I believe that we need to adopt a practice-oriented approach to understand the place of computational cognitive models, and specifically the ways they contribute to understanding of human cognition.

Since computational cognitive modeling consists of various paradigms, and there may possibly be various aims of modeling within one paradigm (e.g. simulating an experimental study and fitting to the experimental data, testing a hypothesis, providing a description or explanation for some phenomena, etc.), an understanding of the *whole* practice may in principle be provided by finding out the invariant features of all modeling paradigms and carrying out the discussion on these features, and possibly doing it by choosing case studies that are representative of the particular modeling paradigms. Because of time limitations, this thesis restricts itself with one of the paradigms of computational cognitive modeling, namely modeling with Soar cognitive architecture (Laird, Newell & Rosenbloom, 1987), and attempts to analyze that particular modeling paradigm with a case study, namely Q-Soar (Simon, Newell & Klahr, 1991) which is representative of a class of models among all Soar models. That class of models is the one with two primary characteristics; firstly, the modelers do not just give the model, but provide the construction process of the model, and secondly, the computational model is more than a pure simulation of the experiment which has been made to study the particular cognitive behavior for which the model is built. As such this thesis can be considered to serve as a *part* of the big project of providing an understanding of the whole practice of computational cognitive modeling.

The structure of the study is as follows. In Chapter 2, the major views about the nature and use of models in philosophy of science is outlined to show the change of attitude towards models historically, and to lay the groundwork for choosing Morgan and Morrison’s (1999) account to make use of in the analysis of modeling with Soar cognitive architecture in this thesis. In Chapter 3, first, the enterprise of computational modeling in cognitive science is addressed with an emphasis on symbol system approaches and cognitive architectures. And then an analysis is presented with a case study, Q-Soar (Simon, Newell & Klahr, 1991), in the light of the ideas of Morgan and Morrison (1999) introduced in Chapter 2. Chapter 4 includes conclusions and discussions.

CHAPTER 2

THE CONCEPT OF MODEL IN PHILOSOPHY OF SCIENCE

Although it is possible to date back the concept of “model” to early Greek in philosophical literature, my concern in this thesis focuses on its more disciplinary treatment in the discipline of the philosophy of science whose history is accepted to have begun in the first half of the twentieth century. Thus, in this chapter, I review a century’s work on models in the philosophy of science literature. My aim in doing this is to make use of the ideas on the concept of model for my primary work of providing an understanding of computational models of human cognition in cognitive science research. Therefore, the analysis towards an understanding of computational models is philosophical in character. (At this point it is important to note that by making use of the literature of the philosophy of *science,* I am not strictly assuming that cognitive science as a discipline *is* a science or not. Rather, I intend to look at it as a research area, whether it is qualified as science or not, that produces a kind of knowledge to understand human cognition.)

There have been various arguments about models throughout the history of the philosophy of science. This variety seems to be partly due to the variety of the things in the sciences qualified as model. However, there is a common point in all these treatments, which is their being discussed in relation to theories and/or phenomena. As a result, it can be claimed that models have not attained an independent status especially from scientific theories until recently. When we look at the history of the philosophy of science literature we see three major views on models. Historically the first two views come within the formal accounts given for the nature of *scientific theories*, and the third is a view focusing basically on the concept of *model* itself.

2.1 Received View

The first view, which is generally called the "received view" in the literature, comes with the analysis of scientific theories in the early twentieth century (Campbell, 1920; Carnap, 1939; Hempel, 1965; Nagel, 1979). At those times, the main focus of the philosophers of science seems to be the nature of the scientific theories which are supposed to teach us about the world. In these accounts, scientific theories are taken as static objects and a general formal account is given regarding their nature. Specifically, scientific theories are formalized as to be axiomatic systems (as in logic). Basic laws of the theory are considered to be axioms of the axiomatic system, and all the other knowledge that the theory contains is supposed to be derived (by means of logical deduction) from the axioms of the theory. Models in this account are an *interpretation* of the logical structure of the theory as in the *model theory* of logic.

There have been differences in the issue of the role of models in science among the philosophers who incorporate models in the formal accounts that they give for the nature of scientific theories. According to Carnap (1939), models may only have an aesthetic, didactic, and heuristic value for understanding, but cannot be necessary components of scientific practice.

Another philosopher who sees models as not essential components of science is Hempel. Nevertheless, he presents the ways models may be useful in science mainly considering the scientific explanation they provide (Hempel, 1965).

In his account, theoretical models in social sciences, particularly the ones in psychology, sociology and economics have the status of theories but with limited scope of application. Particularly, as the basic assumptions, they state interdependencies between various factors (i.e. *variables* or *parameters* in the respective fields' sense) which may be observable or measurable, or theoretical and empirically interpreted. As their basic characteristics are different from advanced physical theories, the class of phenomena that they are concerned with is quite limited. Moreover, the formulation of interrelations within this limited domain may be oversimplified. And additionally, some factors may be disregarded despite the fact that they are more or less relevant. Hempel's explanation concerning the way these models provide explanations is by means of inference from basic hypotheses embedded in them, concerning empirical phenomena. As such, models can be put into tests for explanation and prediction. The nature of the explanation they provide is another issue, which will not be discussed here[1]. However, he notes that the limited scope of application and an approximate validity in this limited scope may restrict the explanatory and predictive value of theoretical models.

Campbell (1920) and Nagel (1979) have a different position in their treatment of models among the previous philosophers although they incorporate models in a formal account given for the nature of scientific theories. Their common point is that they see models as essential components of scientific practice. In Nagel's account, scientific theories are identified with the following three components:

(i) An abstract calculus that is the logical skeleton of the explanatory system in which basic notions of the system are implicitly defined,

(ii) A set of rules that relates the concrete materials of observation and experiment to the abstract calculus,

(iii) A model,

---

[1] The literature of philosophy of science provides several accounts for the explanations given within sciences. An attempt to view the type of explanation given by computational models in the light of those accounts would be another thesis work, and considered to be out of the scope of this thesis.

where model is considered as an interpretation for the abstract calculus in terms of more or less familiar conceptual or visualizable materials.

It is important to note that these components are identified for the purposes of analysis, and are not claimed to be explicitly stated by scientists when they give theories, or to provide the order that scientists follow in the construction of scientific theories. As such, he construes models as one of the essential components of scientific theories.

With his analysis, Nagel (1979) distinguishes several roles models play that makes them essential components of scientific practice. First of all, they may suggest ways of expanding the theory by raising questions on the already constructed theory for further research. Secondly, they may serve in relating the theoretical concepts to experimental concepts in theories. Nagel explicitly states that models are *not* the things that actually relate these two types of concepts but they may suggest which theoretical concepts can be associated with experimental concepts. For example, the model for the kinetic theory of gases suggest that the theoretical expression "change in the total momentum of the molecules striking a unit surface" can be associated with the experimental concept of pressure.

Other than the heuristic value of models in the construction and use of theories, Nagel also acknowledges that they contribute to the establishment of systems of explanations. This role is attributed to them since a newly constructed theory is considered as not only assimilating to the already familiar but may in some way be an extension or generalization of the laws and theories which are assumed to hold for the model itself that initially had a limited scope.

However, it should be noted that the use of models in scientific practice has some potential dangers as Nagel (1979) states. First of all, it may be possible that some irrelevant feature of a model is attributed to the newly constructed theory so that the theory is routed in unprofitable directions. Secondly, the model may be confused with the theory itself, which would be the extreme case of the first danger. Despite being aware of these dangers, one cannot decide in advance whether a model will be heuristically valuable or not since it is usually only possible after its use (Nagel, 1979).

Although not classified together with the previous philosophers in the philosophical literature, Duhem (1954) needs to be mentioned for his criticisms for the use of models in science. He states that since models lack logical structure, they would bring complication to the minds that seek logical linkage among all the parts of a theory. This is so because Duhem thinks that for the scientists that use models, a theory is a model of some group of experimental laws for the pleasure of imagination rather than for satisfying reason. And since it is possible to construct other models for other groups of experimental laws, and it is permissible to have one experimental law included in multiple models, this would not bring the order and simplicity that would be given by a theory identified as having a logical structure. Nevertheless, Duhem does not fully deny the fruitfulness of the use of mechanical models in discoveries on the basis that there are no rules that characterize discoveries, and models may also guide discoveries. But, he attributed little role to them compared to abstract theories in the progress of physics.

2.2 Semantic View

This logical treatment of scientific theories and the view that models are interpretations of the logical structure of theories continue to exist till 1960s. Beginning in the 1960s, counter arguments was developed criticizing the former view for not providing an adequate understanding of the nature of the scientific theories, basically because it *identifies* theories with their linguistic formulations (Suppe, 1977).

One such account is given by Suppe (1977, 1989) for physical theories. His point is that identifying theories with their linguistic formulations- that is, considering that the theory *is* its linguistic formulation- will reveal a distorted picture of the nature of scientific theories for several reasons. First of all, a theory may be given a number of alternative linguistic formulations. (One such example is classical particle mechanics which is sometimes given a Lagrangian formulation, and sometimes a Hamiltonian formulation). Second, a linguistic formulation of a theory may not reflect the characteristic features of a theory. Third, examining the characteristic features of *formulations of theories* may lead one to attribute those features to the *theory* itself, which actually does not possess. On the basis of these considerations, Suppe suggests to examine theories directly, rather than to examine their linguistic formulations and as a result to give formal accounts (e.g. formalizing them as axiomatic systems) to understand their nature.

In the analysis of Suppe, linguistic formulations of theories are considered to provide *descriptions* of the theories, whereas the theory itself serves as a *model* for its formulations. In other words, theories are thought to be abstract structures qualifying as models of their linguistic formulations.

This new account, generally known as the "semantic view" in the literature, provides a formal account, too, but this time the tool used to formalize the nature of the scientific theories is not the axiomatic system. Rather, the nature of theories is identified with mathematical structures –the kind of structure changes depending on the philosopher (Suppe (1989), van Fraassen (1980), Giere (1988)) - or more informally with models. Thus, this view of the nature of the theories does not distinguish theories and models as distinct entities, but rather construes theories as a family of models (Giere, 1988).

2.3 Contemporary View

In both of the earlier accounts, the attitude for understanding science is a formal static one, focusing on the nature of the essential *components* of the scientific practice, namely, theories and models. Beginning in the 1980s it seems to have been established that for understanding science, one needs to look at the *practice* or the *activity* scientists are involved in. That is, the argument is that the philosophy of science should consider the dynamics of scientific practice. This necessitates not just looking at the already existing components of science, but their construction, development, and use. When one has such an attitude, models and the activity of modeling seem to deserve a distinct treatment from a philosophical point of view compared to the earlier *theory-focused* analyses in the philosophy of science. Such an argument has been most extensively given by Morgan and Morrison (1999). Considering a wide range of models in both natural and social sciences, they have provided a framework in which models are treated with four aspects, namely with their construction, their use, the representation they provide, and things we learn by building and using them.

Morgan and Morrison's (1999) basic criticism to the earlier views on models in philosophical literature is that earlier accounts always characterize models as a subsidiary to some background theory, and they state why they differ from them with the following words:

> *Viewing models strictly in terms of their relationship to the theory draws our attention away from the processes of constructing models and manipulating them, both of which are crucial in gaining information about the world, theories and the model itself* (Morgan & Morrison, 1999; p.10).

Their suggestion is to try to see the *autonomy* of models, and so to understand the dynamics of modeling and its impact on the broader context of science.

The main point in the historical movement is that we move from an objectivist position to a more cognitive position in which, to understand scientific activity, we not just look at the nature of its products, but the practice itself, considering the issue of the process of construction and use, what we learn from them, and what they represent. It is important to note that although this last account differs from the earlier accounts in focusing on the model itself, it incorporates some of the ideas about the use of models in science given by the earlier accounts, especially by the ones that see models as essential components of science. The analysis provided in this thesis work takes this last view as its reference for the reasons explained in the beginning of Chapter 3. Therefore, that account of models is presented again in detail in the analysis part, in Chapter 3.

CHAPTER 3

COMPUTATIONAL MODELING IN COGNITIVE SCIENCE

3.1 What is Computational Modeling?

Computational approach in cognitive science takes its roots from the field of artificial intelligence, a branch of computer science, whose subject matter is intelligence. Two leading figures in artificial intelligence research, Newell and Simon (1976), have laid out the ability to store and manipulate symbols as the structural requirements for intelligence. This requirement is characterized in their *Physical Symbol Systems Hypothesis* which states that, *"A physical symbol system has the necessary and sufficient means for general intelligent action."* (Newell & Simon, 1976; p. 116), in which a physical symbol system is defined to consist of a set of physical patterns called symbols, and a collection of processes that operate on these symbols to produce symbol structures. The hypothesis can be said to lay the groundwork for the computational modeling of human cognitive phenomena. Although the concentration of the hypothesis is not restricted to human intelligent behavior, the best-known intelligent system is defined as being human by the researchers. Thus, computational modeling of human behavior with the adoption of the symbol system approach has been developed as a methodology of studying cognitive behavior integrating sources from psychological observations and experiments.

3.2 Approaches in Computational Modeling

There have been other paradigms within the computational approach to studying cognitive behavior in cognitive science, connectionism being one of the most extensively studied in a variety of cognitive domains (Rumelhart, McClelland, 1986). The main idea in connectionist systems is the parallel-processing of sub-symbols using statistical properties, and is inspired from the known neurophysiology of the brain.

Another approach in computational modeling of cognition is the dynamical system theory. A dynamic system is defined to be a system, which changes over time according to some lawful rule. The goal of these systems is to characterize the kinds of changes in the system with mathematical formalisms (Elman, 1998). The attempt within cognitive science with this approach is to model cognition considering it as a dynamical system.

Besides the computational work within specific cognitive domains adopting a particular computational approach (symbolic, connectionist, or dynamic), there have also been attempts to give domain-generic computational models on which a range of different cognitive behaviors can be simulated. These are architectural approaches to cognition (Anderson & Lebiere, 1998; Newell, 1990; Kieras, 2004). Architecture is a term mostly referred in computer science to describe the hardware organization of computers chosen by the manufacturer. It specifies a set of fixed structures and mechanisms that allow a variety of software to run on it, independent of the software. In a similar fashion, a cognitive architecture is considered to specify common structures and mechanisms that are fixed underlying all human cognitive phenomena (Lehman, Laird, Rosenbloom, 2006). As such, a cognitive architecture is given as a theory of what is common to all cognitive behavior. Providing domain content of a particular cognitive behavior into the architecture, a model of that particular cognitive behavior is given.

The idea of architecture has been a key element for some researchers who tend to see cognition as a big picture (Lehman, Laird, Rosenbloom, 1996). Viewing a theory in any of the disciplines that is said to study human cognition as taking a *portion* of this big picture, the concern of these people is stated as:

> *When each discipline throws its set of pieces out on the table, how do we know that there is any set of pieces that will allow us to recover the big picture?* (Lehman, Laird, Rosenbloom, 1996).

Gathering all the pieces into the whole picture by means of presenting "unified theories of cognition" is the suggested answer to this question given by Allen Newell (1990). SOAR (Laird, Newell, Rosenbloom, 1987) is one of those cognitive architectures, which is developed for this ultimate aim. Modeling with SOAR is the practice that is analyzed in this thesis, as a part of the big project of providing an understanding of the whole computational cognitive modeling practice. At this point, it is important to note that choosing the paradigm of modeling with SOAR does not carry the assumption that modeling with SOAR is a representative paradigm of computational modeling. It attempts to be a piece of the big project mentioned above by providing the details of this particular practice, and to take part for determining the invariant features of computational cognitive modeling practice with the examination of other paradigms in pursuit of giving general account of computational cognitive modeling as a methodology employed in cognitive science.

3.3 An Analysis of Computational Models: An Exemplar Model with Soar

Considering computational modeling as a distinct and widespread methodology in cognitive science research, I believe Morgan and Morrison's (1999) framework is an appropriate one to present an understanding of the modeling practice in cognitive science since for one thing it treats modeling as a distinct practice from the theory-centered accounts which favor the theory to the model in understanding scientific practice, and focuses on the modeling practice. For another thing, within the account of modeling, it focuses on the *dynamics* of modeling, rather than taking models as static objects of science. The impact of this point is that the account given may be of interest of scientists as well as philosophers since it may shed light on what scientists actually do in their scientific activity and make them think about their own practice. Considering the presence of computational cognitive models that are just given without providing the construction process, or that are just given to fit some set of empirical data, and as a result coming up with the question of whether there is anything special provided by computational cognitive modeling apart from being a tool for simulation of experimental studies, I appealed to Morgan and Morrison's (1999) framework since the emphasis on dynamics of modeling has the potential to analyze the practice apart from what is provided by the modelers. For this aim, I considered one by one the four elements discussed in the framework, which are the construction and development of models, the function of models, the representation they provide, and things we learn from models in the following sections. However, before that, let me introduce the basics of Soar architecture.

3.3.1 Basics of Soar

Soar is characterized to be a cognitive architecture. So, first of all, let's begin with defining what a *cognitive architecture* is. A cognitive architecture is defined to be a theory of the fixed mechanisms and structures that underlie human cognition (Lehman, Laird, & Rosenbloom, 2006). The idea of cognitive architecture stems from seeing cognition as one big picture and trying to build *unified theories of cognition* that gather all separate pieces of theories that account for various regularities we observe in human cognition and factor out what is common across all cognitive behavior. Figure 3.1 below is an abstract representation of the Soar architecture with its main elements:

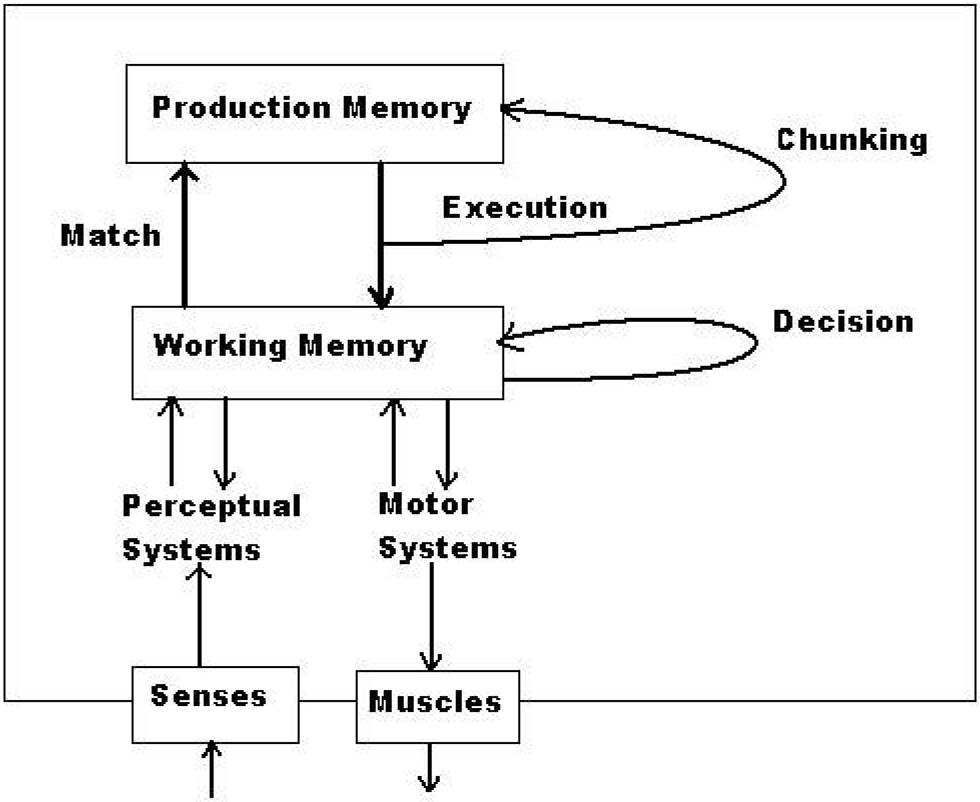

Figure 3.1. The overview of Soar architecture (adapted from Taatgen (1999)).

The most crucial thing regarding modeling with cognitive architectures is that architecture by itself does nothing; it just provides a fixed set of mechanisms and structures to underlie any cognitive behavior. In order to produce behavior, it needs *content*, which is the theory of the domain-specific knowledge for the particular cognitive behavior. This is clearly stated by

BEHAVIOR = ARCHITECTURE + CONTENT (Lehman et. al., 2006).

The foundational idea in representing behavior through time in Soar is *problem space* which dates back to the earliest days of artificial intelligence research. The idea is to view behavior as a space of decisions made over time. It is abstractly represented by a triangle including various elements which will be introduced soon (Figure 3.2).

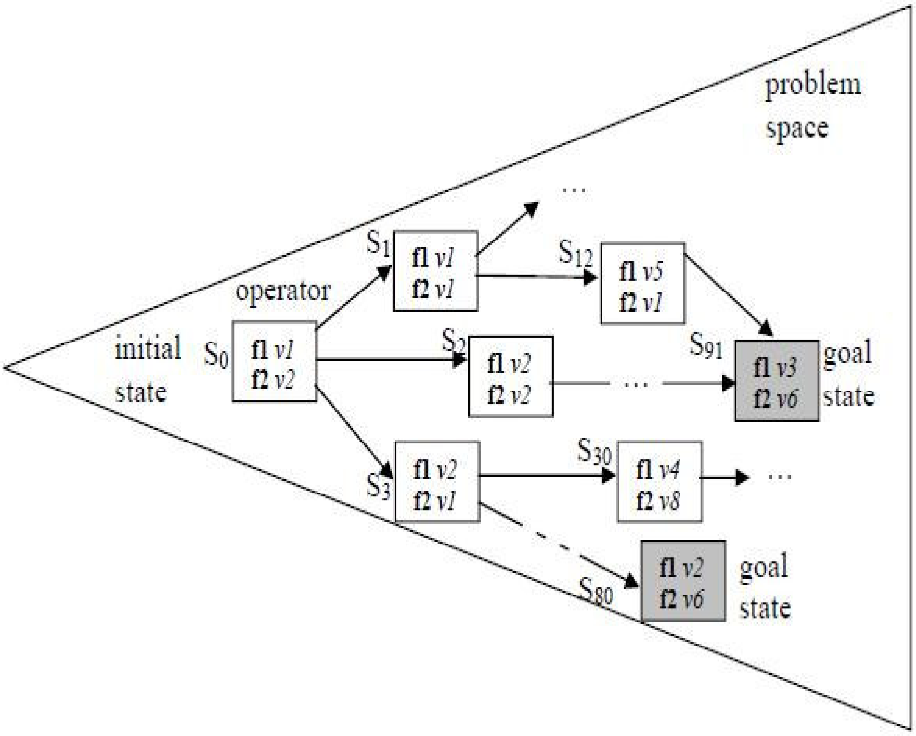

Figure 3.2. Abstract form of the idea of problem space (adapted from Lehman et. al. (2006)).

With this idea, any situation of a person at any moment is considered to be a *state*, and behavior is thought to be as movement through these set of states by means of *operators* for accomplishment of *goals*. Specifically, a state is described by a selected set of features (**f1**, **f2** in Figure 3.2) and their possible values (*v1*, *v2*, etc. in Figure 3.2) to represent any situation. Any change in the situation means a change in the state, and this change is represented by means of applying *operators* to the current state to arrive at a new state. Thus, movement through the problem space is accomplished by applying operators. It is important to note that the selection of an operator among many alternative ones at any moment is guided by *the principle of rationality*, which is defined as: "If an agent has knowledge that an operator application will lead to one of its goals then the agent will select that operator" (Lehman et. al., 2006). Altogether, the elements of goals, states, operators, and problem spaces are considered to correspond to different kinds of knowledge a person has.

In order to see how the idea of viewing behavior as movement through a problem space works, we need to consider a sample case and try to see how it is formulated as a problem space. This case is a simple one in which there are three blocks on a table named A, B and C, and the goal is to put them in a vertical arrangement where block A is on top of block B, block B is on top of block C, and block C is on top of the table. Figure 3.3 below displays the graphical representation of the current scenario:

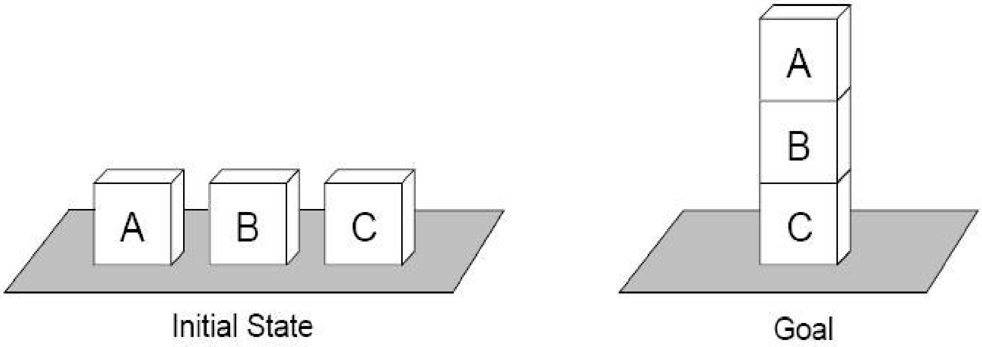

Figure 3.3. Graphical representation of the sample scenario (adapted from Soar's User Manual).

Looking at the situations described in the scenario (starting from the initial state and arriving at the goal) as a space of decisions over time, one can obtain the following problem space description of the scenario (Figure 3.4):

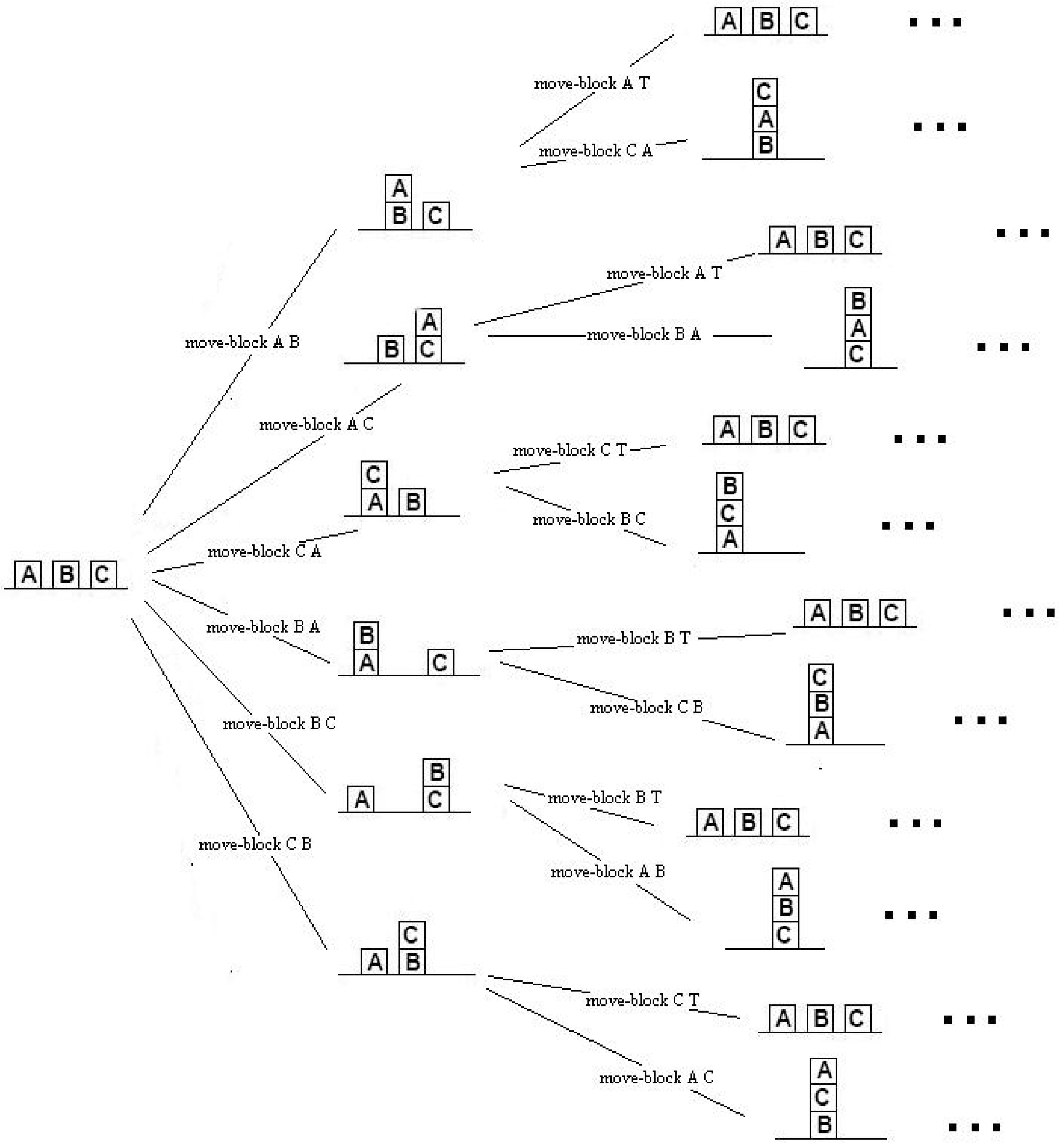

Figure 3.4. A space of decision made over time for the scenario (adapted from Soar's User Manual)

The figure above resembles a problem space, but in order to formulate it fully and formally, other elements within the problem space, namely the states, operators, and goals, must be specified. The representation of an exemplar state (initial state in Figure 3.3) is as follows (Figure 3.5):

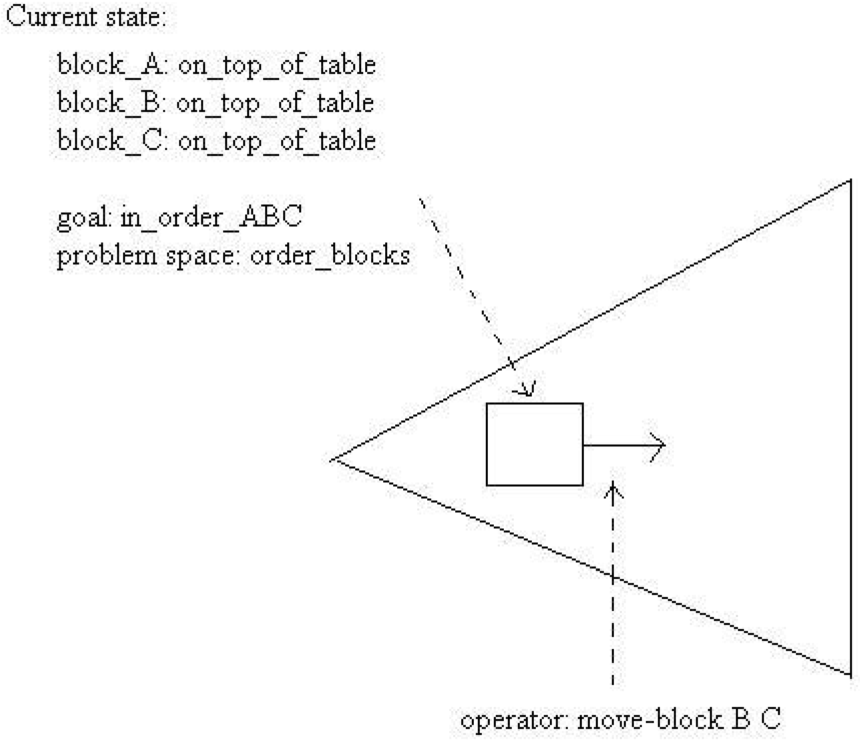

Figure 3.5: The representation of state, goal, problem space and operator for the blocks scenario.

So far, we have described the basic elements to represent different kinds of knowledge in Soar. Let's now describe the architectural processes together with the details of the structures for representing knowledge. There are basically two kinds of *memory* structures to represent knowledge in Soar: *Long-term memory (LTM)* and *working memory (WM)*. Long-term memory is a structure holding knowledge that exists independent of the current situation. It is represented as a *production system* and so knowledge is represented by *if-then rules*. On the other hand, working memory holds knowledge that is relevant to the current situation. It is represented as a set of features and values that make up the current state (and substates), which might include representations of the current problem space, operator, and goal. These elements arise either through perception or through retrievals from long-term memory.

The relationship between these two memory structures is part of the main architectural process, called *decision cycle*, for representing cognitive behavior. At any moment, the “if” portion of each rule in LTM is tested for whether there is a match with the elements in WM. If there is a match, the rule is said to be *triggered*. This means that the “then” portion of the particular rule(s) is applied either by suggesting changes to the goal, state, and operator, or by sending messages to the motor system. The figure above (Figure 3.1) illustrates the relationship between the two memory structures together with the perception-motor interface.

The *decision cycle* is the basic architectural process consisting of three phases; elaboration, decision, and application. During elaboration phase, rules in the LTM are accessed in parallel to suggest new operators or to evaluate operators. In the following decision phase the suggested operators in the previous phase are evaluated. In the application phase, rules fire to modify the state. There are several other architectural mechanisms. These can be classified under two headings; impasse signaling and learning mechanisms.

Impasse signaling is an automatic mechanism to signal a lack of knowledge for the decision procedure to select an operator. When an impasse occurs, a new substate is automatically created whose goal is to resolve the impasse. On the other hand, there are four architectural learning mechanisms, including chunking, reinforcement learning, episodic learning, and semantic learning.

After this basic introduction to the structures and mechanisms of Soar architecture, let's briefly describe the "model building" process within this architecture. The modeler begins by specifying the domain knowledge the person needs for the particular cognitive behavior. That domain knowledge comes either from theories or empirical data or both. This step includes describing the relevant goals that drive the behavior and knowledge to accomplish them. Once there is this informal description (called knowledge-level description), knowledge can be tied to goals, states and operators of the Soar architecture. And the next step is to specify the relationships between different levels by the potential impasses (the hierarchical representation of problem spaces for the particular cognitive behavior) and the kinds of knowledge that will be missing in accomplishing the particular cognitive task at hand.

3.3.2 Analysis of Modeling with Soar

Construction and Development

The basic claim of Morgan and Morrison's (1999) account of models is that models are *autonomous agents*, and as such they function as *instruments* of investigation in understanding both *theories* and *phenomena*. To see that this claim is true, one feature they have looked is the construction of models. They state that although it is commonly thought that models are entirely derived either from theory or from data, a closer look at the way models are constructed will show the sources of their independence, since they involve elements both from theory and from data, and additionally involve "outside" elements (Morgan & Morrison, 1999).

Morgan and Morrison (1999) point out that no general rules for model building have been provided in scientific practice or in philosophical analysis so far, and state that it has been generally considered to be an art. The account they provide characterizes model construction as a process of choosing and integrating a set of items that are considered to be relevant for the task at hand. These items are typically elements from theories, from empirical data, and anything the *modeler* provides (mathematical formalisms, metaphors, stories) which constitute the "outside" elements. The integration of those distinct elements into a formal system provides the key relationships between numbers of variables. The presence of the outside elements in the model, and the fact that theory by itself does not determine the model form sets the model as partially independent both from theory and data. And this feature of models, that is their partial independence, enables them to offer new theoretical understanding of the phenomena under investigation.

One important point made for distinguishing models as separate elements from theories and data in this account is that even without the process of integrating a set of items from distinct sources, models still have the status of partial independence since they always involve certain simplifications and approximations compared to the theoretical requirements and/or data conditions (Morgan & Morrison, 1999; pp. 38-65). On the other hand, it is also argued that since we expect to use models to learn about theories and phenomena, models must in some way be related to theory and phenomena.

Now let's look at the construction of models of cognition using Soar as a paradigmatic case of computational modeling practice in cognitive science research. First of all, it is very important to note that the construction of the architecture itself (without any domain content) includes elements both from scientific theories and observational data, and additionally, ideas from artificial intelligence research. The list of multiple constraints that shape mind provided by one of the creators of Soar, Newell (1990), as the considerations in the construction of the Soar architecture exhibits the elements coming from scientific theories and observational data. The list that presents the constraints on an agent possessing mind is as follows:

1. *Behave flexibly as a function of the environment*
2. *Exhibit adaptive (rational, goal-oriented) behavior*
3. *Operate in real-time*
4. *Operate in a rich, complex, detailed environment*
    a. *Perceive an immense amount of changing detail*
    b. *Use vast amounts of knowledge*
    c. *Control a motor system of many degrees of freedom*
5. *Use symbols and abstractions*
6. *Use language, both natural and artificial*
7. *Learn from the environment and from experience*
8. *Acquire capabilities through development*
9. *Operate autonomously, but within a social community*
10. *Be self-aware and have a sense of self*
11. *Be realizable as a neural system*
12. *Be constructible by an embryological growth process*
13. *Arise through evolution* (Newell, 1990).

The elements coming from artificial intelligence research show themselves in the structures and mechanisms of the full architecture, which are described in the previous section (e.g. problem space, production system). Thus, any model building process within Soar architecture carries with it the elements involved in the construction of the *architecture*. These elements might be interesting from the point of view of cognitive science if they carry assumptions about cognitive processes rather than be solely computational structures that enable the production of some set of output given some set of input. In fact, the structure *working memory* in the architecture can be considered to correspond to the concept with the same name, working memory, in the psychological literature, which is defined to hold knowledge for a short period of time. In addition, the access of long-term memory by means of some knowledge in working memory as an architectural process resembles the associational nature of human memory. That is, we as humans normally retrieve memories from the past as long as they are associated in some way to the knowledge we possess at a moment.

After pointing this out, let's look more closely at the *model building* process within Soar architecture. To discuss it in more concrete terms after this point I will introduce a model developed with Soar architecture, called Q-Soar, by Simon, Newell & Klahr (1991). This model is an example of a class of models which does not aim purely to provide a simulation of a specific experimental study and fit to some set of empirical data, but rather aims to give a computational description or explanation to some particular phenomena. This feature of the model is the reason of its selection as a case study to carry out an analysis on. Because, I think this kind of application of computational cognitive modeling distinguishes it from experimental studies and makes it an alternative for other ways of providing description or explanation of some phenomena.

Q-Soar is a computational model of a phenomenon that has been studied in developmental psychology, which is the learning of number-conservation knowledge in children. The creators of the model refer to Piaget's definition (as cited in Simon, Newell & Klahr, 1991) to describe the phenomenon:

> *We call "conservation" (and this is generally accepted) the invariance of a characteristic despite transformations of the object or of a collection of objects possessing this characteristic. Concerning number, a collection of objects "conserves" its number when the shape or disposition of the collection is modified, or when it is partitioned into subsets* (p. 1361).

To model the phenomenon, the authors chose a particular experimental study from the literature, which is the study of Gelman (1982) for its having well-defined procedures and clear quantitative outcomes.

Very basically, the experimental study is as follows. Children of three and four year olds were first trained in a brief session using small collections of objects (n = 3-4) in both equivalence (two rows of equal number) and inequivalence (two rows of unequal number) relations, and then tested with both small (n= 4-5) and large (n = 8-10) collections. In the experiment, one experimental group and two control groups were used (Simon, Newell & Klahr, 1991). The experimental group was trained with the following training material in Figure 3.6 and the following steps:

(1) The display was presented in one-to-one correspondence and the child was instructed to count the number of items in one of the rows.
(2) That row was covered by the experimenter and the child was asked, "how many are under my hand?"
(3) The child was instructed to count the number of items in the other row.
(4) That row was covered by the experimenter and the child was asked, "how many are under my hands?"
(5) The child was asked to judge whether the two uncovered rows contained "the same or different number" of items.
(6) While the child watched, the length of one of the rows was spread or compressed.
(7) The experimenter pointed to the altered (or unaltered) row and asked, "are there still n here?"
(8) The experimenter pointed to the other row and asked the same question.
(9) The child was asked whether the pair of rows had the same number or a different number of items, and to explain his/her judgment.

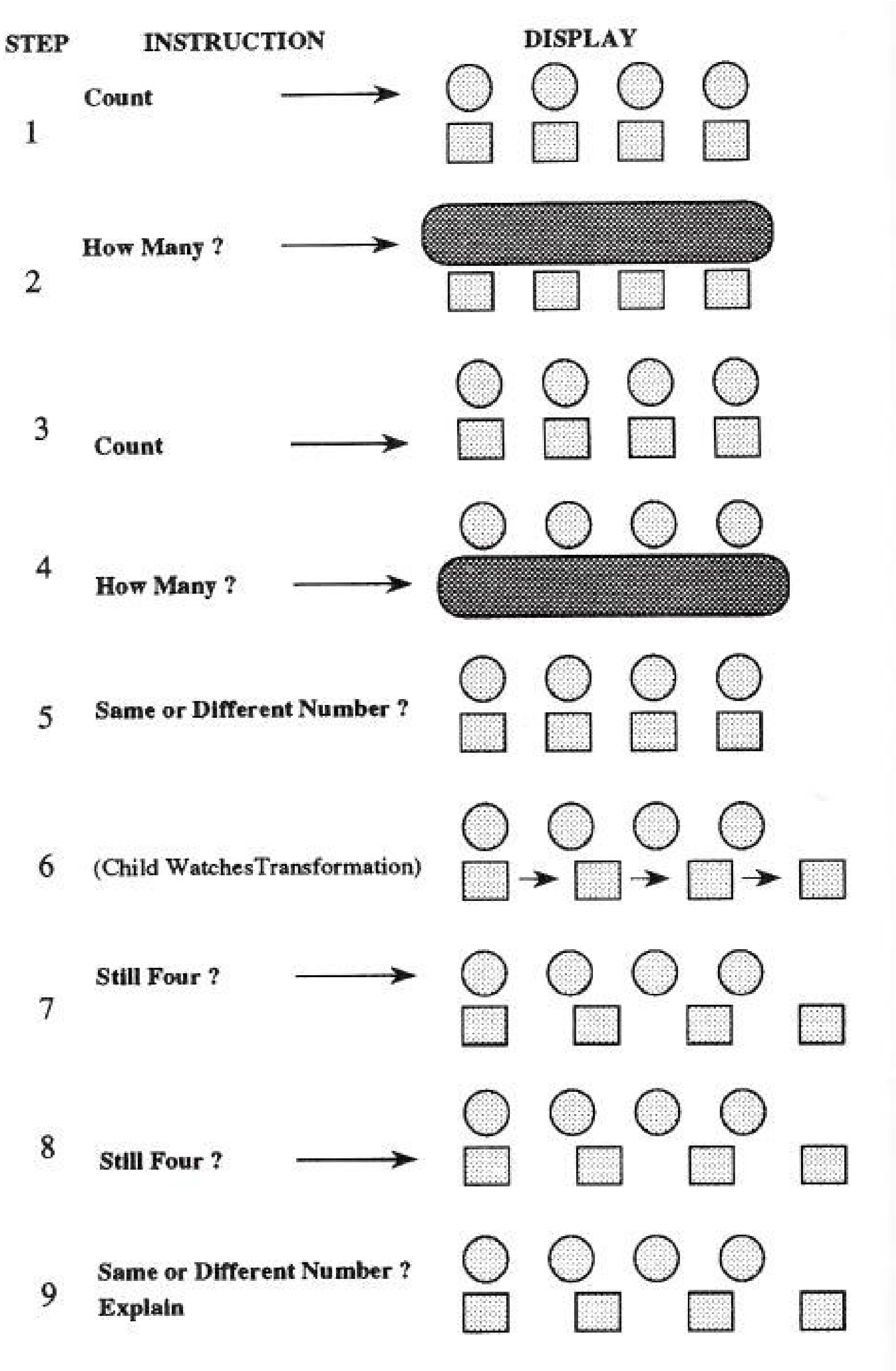

Figure 3.6. Graphical representation of the experimental procedure (adapted from Simon, Newell & Klahr (1991)).

The training in the control groups (One is called *Cardinal-Once* group and the other is called *No-Cardinal* group) were as follows: Children in the Cardinal-Once group were exposed to only one row (of three or four items). For that one row, they were exposed to steps 1-2 and 6-7 listed above. The other control group, No-Cardinal, simply counted single rows of three or four items, and were exposed to step 1 listed above.

After the experimental or other procedures, conservation tests were applied. Each child was given four different tasks (large or small set size, and equal or unequal numbers of items in the two rows). Small sets included either 4-5 or 5-5 items and large sets included either 8-10 or 10-10 items. Children are discouraged from counting. The results show that the Experimental group passed the tests at a rate of 70% (3 and 4 years olds did equally well), while the untrained group, No-Cardinal group, had a pass-rate of ranging between 0-15% for all 3 and 4 years olds. There was difference between 3 and 4 years old in the Cardinal-Once group with pass rates of 9% and 46%, respectively.

Evaluating the results of the experimental study (specifically looking at the difference between Experimental group and Cardinal-Once group), the model builders come up with their general hypothesis about the knowledge and strategies used by the children in the context of the conservation experiments, calling it Q-Theory (to distinguish it from the model Q-Soar).

Q-Theory states that the strategies used by the experimental group are different than the ones used by the Cardinal-Once group as a result of different training sessions. To model this piece of knowledge, the authors state the basic things the system must be able to specify as follows:

> *(1) The knowledge state prior to the training (i.e. a non-conserving child).*
> *(2) The encoding of the collection(s) prior to transformation. This will include the salient features such as number, length, and density, as well as other features that may ultimately be irrelevant for the task at hand.*
> *(3) The encoding of the relation between collections (for the experimental group)*
> *(4)The encoding of the collections(s) following transformations*
> *(5)The encoding of the physical aspects of the transformation (e.g. salient motion, how objects were moved, how many were moved, direction of movement); and*
> *(6)New knowledge acquired from repeated trials of the kind presented to both the experimental group and the Cardinal-Once group*
> (Simon, Newell & Klahr, 1991).

The model Q-Soar is constructed conjoining Q-Theory with the assumptions of Soar as a cognitive architecture (specifying the problem spaces). The creators of the model specify their design assumptions in building the model with the following words:

> *Having been shown a transformation to a set of objects, the child first categorizes the transformation and then initiates a conservation judgment about the transformation's effect. Ideally, categorization will identify the observed transformation as an instance of a larger class, with effects that are known to be associated (through chunking) with this class. If not, then pre- and post-transformation values created by measurement processes are compared to determine the effect of the transformation. The learning over this processing creates new knowledge about this kind of transformation, which will become available on future occurrences in similar contexts* (Simon, Newell & Klahr, 1991, p. 1375).

Consequently, from what the modelers described in the construction of Q-Soar, we can conclude that the model building process in the first instance involves the interpretation of empirical results, consideration of theoretical assumptions, and production of general hypotheses. Following it is the specifications with the architectural structures and mechanisms (problem spaces, states, goals, operators, memory elements). In terms of the framework of Morgan & Morrison (1999), the model is said to be autonomous since it is not directly derived from theory or empirical data. And this feature gives it the potential to provide new theoretical understanding of the phenomena. This is so since the authors use the model to anticipate about the necessary steps toward a full theory of conservation in other domains (other than *number*). For this aim, they evaluate the model in terms of Piaget's (1964) several criteria for conservation. In the evaluation, they use one of the learning mechanisms of Soar, chunking, as their framework to make theoretical claims about some issues regarding children's learning of the number-conservation (Simon, Newell & Klahr, 1991).

Function

According to Morgan and Morrison (1999), one feature of models that give them their autonomous status in scientific practice is their ability to function in a variety of ways. One of the common uses of models is to aid in the construction and further development of theories. Based on the examples from science that they study in their analysis, they specifically point out the following ways models function in theory construction and development:

- To explore certain characteristics of a theory
- To explore processes for which theories do not give good accounts
- To function directly in experimental studies as an instrument (Morgan & Morrison, 1999).

In the light of these possible ways of functioning of models in science, let's now see the ways a computational model with Soar cognitive architecture can function. It is important to note that it is not always possible to see all the ways models function just in one computational model. As it is the case for models in science in general, different computational models may function in different ways.

Recall the computational model introduced in the last section, the model of children's learning about number conservation. As a computational model simulating the experimental study of Gelman (1982), it can be considered to function as an instrument in experimental studies. Although not mentioned in the specification of the particular model, such uses can occur in the lack of immediate resources for actual experimental studies (e.g. subjects). It is important to note that this function of computational models, that is, their being simulation tools, is a feature that can be generalized to other models built with Soar and also to models with other computational paradigms that can provide the simulation of some experimental study. Moreover, new hypotheses which are not tested or capable of being tested in real experimental conditions can be tested with the computational model. Thus, it both plays role in exploring certain characteristics of the theory by means of manipulation to the experimental conditions within the computational model, and has the potential to function in further development of the theory (for example to expand the theory to account for conservation of things other than *number*).

Additionally, Q-Soar provides a new theoretical understanding of the phenomenon by providing a new framework, namely a computational description, to the children's learning behavior in the particular task of the experimental study. This gives the opportunity to explore processes for which theories do not give good accounts. In supporting this line of thought, the authors state their aims in creating the particular computational model as to provide a *complete* specification of a set of operations and their interaction with a specified learning mechanism, which has not been given so far by some other kinds of efforts, and to show that it produces the same pattern of behavior observed in children acquiring conservation knowledge (Simon, Newell, & Klahr, 1991; p. 1362).

It is implied that the particular computational model is created and used both to explore the characteristics of the already available accounts and also to explore processes for which available theories do not give good accounts. As for the other models built with Soar, such a function of a model is possible if the modeler does not just aim to provide a fit to empirical data, but actually takes the specifications in terms of architectural mechanisms to account for the underlying mechanisms of the particular cognitive behavior under investigation.

Representation

The representation that models provide is one of the issues discussed in the framework of Morgan and Morrison (1999) that aims to present an understanding of models in science. The understanding of the notion of "representing" in this account is not restricted to the cases, where the model has a mirroring relationship with the phenomenon, system, or theory. They define it as follows:

> *… a representation is seen as a kind of rendering- a partial representation that either abstracts from, or translates into another form, the real nature of the system or theory; or one that is capable of embodying only a portion of a system* (Morgan & Morrison, 1999; p. 27).

Specifically, they discuss the issue of representation provided by the models by *relating* them either to theory or to empirical data or to both. That is, in their terms, a model represents either some behavior described by a theory, or the behavior of the real system being modeled.

In the case of computational models in general and models within Soar in particular, both of these alternatives are possible. For computational models which are built following the principles of an available well-established theory, the model can be said to represent the behavior described by the theory. On the other hand, there have also been computational models, which are built in order to account for some phenomena that have not been given good accounts by available theories. For such cases, the models can be said to represent the behavior of the real systems.

Q-Soar described in the previous section is an example of the kind of computational model that represents both the empirical data by means of simulating the behavior of children, and the theory, Q-Theory which specifies the mechanisms for that behavior formulating them in terms of the architecture's structures and mechanisms. Thus, in addition to the empirical data and the content theory, the model Q-Soar can also be said to represent all the theoretical and empirical elements used in the construction of the pure Soar architecture. For instance, the structure working memory which is defined to hold the current knowledge an agent has in architectural terms can be considered to represent the concept of working memory in the psychological literature which is generally defined to hold knowledge for a short period of time. Furthermore, the architectural process of accessing long-term memory triggered by some elements in the working memory represents a way of remembering the memories of past experiences, which are associated with something we as humans hold as knowledge at that moment.

In addition, there is also a room for *simulations* in Morgan and Morrison's (1999) account. As a result of exploring simulations in physics and economics, they conclude that:

> *… models are capable of representing physical and economic systems at two distinct levels, one that includes the higher level structure that the model itself embodies in an abstract and idealised way and the other, the level of concrete detail through the kinds of simulations that the models enable us to produce* (Morgan & Morrison, 1999; p. 30).

It may be argued that the kind of computational models in physics or economics is different from the ones in cognitive science. Nevertheless, the representation capability at two distinct levels can be applied to computational models in cognitive science as well. Recall that Q-Soar is built to provide structures and mechanisms to account for a phenomenon, which can be thought to be the higher level structure, namely the specification of the content theory in terms of problem spaces, states, and operators. The other level of concrete detail shows itself in the actual simulation of the particular behavior when one tracks it.

Learning

The fourth element in Morgan and Morrison's (1999) account of models is the question of what we learn from models. The simple answer given by the account is that we learn about both theories and the world. Morgan and Morrison (1999) have further investigated how this learning takes place. According to the account, models allow the possibility of learning in two ways; through constructing the model, and using the model.

Constructing a model is characterized as a process of integrating together empirical evidence and theoretical elements, and also some outside elements, as it has been described previously. The process of choosing and fitting available knowledge about the particular phenomenon together is said to be the opportunity to learn both about theories and the world. Considering what the creators of the Q-Soar report about the efforts they gave in the construction of the model, such an opportunity can actually be observed. Those efforts can be listed as follows: The authors firstly review the literature about the particular phenomenon and list the general regularities that have been consistently observed (Gelman's (1982) study together with other studies), which is the means to learn about the phenomenon. They specify their theoretical assumptions, and come up with a general theory (in terms of the kind of mechanisms that can take place) to account for the particular behavior. Then, they conjoin the theory with the Soar architecture, and simulate the behavior of the children in the experimental study. This last work can be said to enable one to think about both the content theory and the empirical data.

Additionally, one can also learn about the Soar architecture in detail by means of specifying the content theory in terms of Soar structures and mechanisms. This opportunity –in fact any modeling activity with the Soar architecture- can make possible to extend the Soar architecture, too, towards a fully-specified unified theory of cognition, and consequently can further development of models of all cognitive domains. A look at the historical development of the Soar architecture will show that among many other reasons, the need to account for a wider range of cognitive behavior is a primary reason to make changes in the architecture (e.g. architectural modifications to model reinforcement learning (Nason & Laird, 2004)).

In the framework of Morgan and Morrison (1999), the other way models allow the possibility of learning is through using them. Using models is characterized as manipulation of models to learn in the first instance about models themselves. Only after learning about and from model's own internal structure one can start to understand the actual systems being modeled. Despite the lack of discussion about some manipulation other than the procedures of the experimental study in Q-Soar, the model, as any computational model, has the potential to be manipulated after its construction by means of testing different sets of data from the original experiment. So, this is another possibility to re-consider the empirical results and the theoretical assumptions. In addition, the model may provide us the opportunity of learning about some associated phenomena other than the children's learning of number-conservation. For instance, this is possible if one attempts to obtain a computational account for the phenomenon of conservation in other domains other than number. Thus, manipulation can take place in at least two ways. In one way, the kind of knowledge provided to the architecture may change (e.g. the details of the experimental study, the empirical data for testing). In another way, the mechanisms may be modified or extended to provide new accounts. As such, it is a model which has the potential to teach us by using it.

CHAPTER 4

CONCLUSION

This study has examined computational modeling practice in cognitive science research taking it as one of the major enterprises that attempts to understand human cognition. Returning to our research question stated at the beginning, it has investigated in what ways computational models contribute to understanding of human cognitive phenomena. Since examination of practices of sciences is a major research area in the philosophy of science, its literature and the ways of examination were thought to be enlightening, and therefore made use of in this study. Among a century-work of ideas in the literature of models, the one that most extensively and most systematically analyzes the practice of modeling was chosen to be the framework for this study, which is the account of Morgan & Morrison (1999). Because, for one thing, it is an account that takes modeling as a primary activity employed in scientific practice so that the focus is on models rather than most-favored components of scientific practice in order to understand science by philosophers of science until recently, namely theories. For another thing, consideration of the *dynamics* of the modeling activity in the framework of this account might be interesting for scientists, as well as philosophers, since it may make scientists to think about their own activities.

Following the way the analysis is carried out in the formulation of Morgan and Morrison's (1999) account, a case study was employed in the analysis to discuss in more concrete terms. This was chosen to be the modeling of a phenomenon with Soar architecture. There is no special reason to choose modeling with Soar architecture among many other paradigms of computational cognitive modeling.

However, at this point, it is important to re-phrase what I have said in the introduction part of this thesis. Any attempt to provide an understanding of computational cognitive modeling in general must firstly figure out the invariant features of various paradigms of computational cognitive modeling, and this might be possible by studying many case studies within each paradigm. Accordingly, the case study that has been used in this thesis work is not claimed to be a representative of all computational cognitive models or of modeling activity in all paradigms. Rather, it is a representative model of a smaller class within models built with Soar architecture, namely models that are not just built to fit to empirical data or be a simulation of some experimental study, but aim to provide a computational description or explanation to the phenomena under investigation as an alternative to other ways of providing descriptions or explanations. As such, what has been provided by the analysis can be generalized to the class of models that this model is representative of. But, this does not mean that one can have implications for other class of models built with Soar in the first instance, and for models built with other modeling paradigms, too.

Let me now summarize what the analysis of the particular case study used in this thesis provides us, and then discuss the implications of these to other models built with Soar and to the computational models built with other modeling paradigms. The analysis of the case study consisted of four parts. The first part examined the construction of the model and its potential for further development after construction. This examination revealed the elements and processes that contributed to build the model. This element of Morgan and Morrison’s (1999) framework, that is the construction, can be considered to be a key element in any attempt to understand any modeling practice in cognitive science since most of the time models of some phenomenon in any paradigm are just given to account for some question under investigation without specifying the construction process. Since the construction process of the case study in this thesis has been provided by the modelers, it is easier to see the elements and ways of integrating the elements to build the model. They are basically theoretical elements, empirical elements and some concepts from AI field (e.g. problem space, production system) which can be qualified as “outside elements” in Morgan and Morrison’s (1999) terms. As for the other models built with Soar or other computational paradigms, even if the construction process is not provided by the modelers, one can attempt to figure out the elements and the process of construction by studying on the model or even by rebuilding the model. In the abstract level, any computational model built with any modeling paradigm may involve theoretical and empirical elements regarding the phenomenon under investigation. On the other hand, the outside elements may vary depending on the assumptions of the particular modeling paradigm.

The second part of the case study explored the function of the model. This examination displays some possible ways this computational model may function. These include the exploration of certain aspects of the theory that existed before the model, providing a new theoretical understanding for the phenomena under investigation with a computational description, and thus forming a framework for further investigation, and being a potential device to function as an instrument for experimental studies. Actually, any computational cognitive model which provides a simulation of some experimental study can function as an alternative to the experimental study. However, computational models whose aims get ahead of fitting to experimental data are the only ones that may take roles in description or explanation of some phenomenon.

The third part examined the question of what the model in the case study represents. The examination showed that the model can be considered to represent several things. These include the representation of the behavior in the phenomenon under investigation, certain aspects of the theory which specified how the phenomenon came to occur, and aspects of the Soar architecture on which the theory and behavior were specified. Studying on any computational model or reconstructing it to figure out the assumptions made by the modelers with available knowledge for a particular phenomenon might help one to understand what is represented by the model as a whole or by certain parts of the model.

The fourth part examined the way we advance our knowledge with the model in the case study. For this aim, the examination considered both the construction of the model and its use after construction. It was concluded that the construction of the model enabled one to learn about the existing regularities observed in behavior; to come up with hypotheses based on these regularities, which advances to gain knowledge by testing the hypotheses; and to perceive the assumptions one had in coming up with the hypotheses, which would enable one to add new assumptions or remove the existing ones in the way of facilitating understanding. Moreover, beside the construction of the model, its use was also considered to be an opportunity for learning. This was stated to be possible with the manipulation of the model in pursuit of certain goals after its construction. To see what may be learned from other computational cognitive models or modeling practices, one may need to involve in the practice, either by constructing models or using them.

As a final note, I want to highlight three points which concern the computational cognitive modeling *in general*: First of all, the analysis carried out in this study shows that the particular computational model supports the main claim about the autonomy of models in Morgan and Morrison's (1999) account. This is so since it is partially independent both from theory and from empirical data as it was illustrated in the four steps of the analysis. That is, the model neither just mimicked the empirical data nor derived directly from the theory. Moreover, its use was shown to differ from the use of the theory by it being both an instrument for simulating the experimental study and a new framework of understanding phenomena and of applying the existing theory. As a result, the learning that took place in the construction and in the use of the model is peculiar to the modeling practice. To see that the claim for the autonomy of models holds for computational cognitive models in general, one must study on the various models in different modeling paradigms. However, in principle one can say that they are solely dependent neither on theories nor on data since they in the first instance involve technical AI concepts, which distinguish them from both theories and experiments.

The second point concerns the applicability of the framework of Morgan and Morrison (1999) to other modeling practices with Soar and other computational cognitive modeling paradigms. With its four elements, namely construction, function, representation and learning, the framework seems to be in principle applied to analyze the other modeling practices. It may be especially useful to evaluate computational models which are just given without specification of the construction process. However, it is possibly not the sole framework to evaluate computational cognitive modeling, but it is the only one that we can benefit from the philosophy of science literature.

As a final point, I want to re-phrase my statement in the introduction part about the place of this thesis work. This thesis can be considered to serve as a *part* of the big project of providing an understanding of the whole practice of computational cognitive modeling. Because, it analyses one case study which is representative of a class of models within one paradigm of computational cognitive modeling, and one must study that kind of representative models of different paradigms in order to figure out some invariant features of computational cognitive modeling in general and to give a general account for the practice.

BIBLIOGRAPHY

Anderson, J. R. & Lebiere, C. (1998). *The atomic components of thought*. Mahwah, NJ: Erlbaum.

Campbell, N.R. (1920). *Physics: The Elements (1920)*. Cambridge: The University Press.

Carnap, R. (1939). *Foundations of logic and mathematics*. Chicago: University of Chicago Press.

Cartwright, N. (1983). *How the laws of physics lie*. Oxford: Clarendon Press.

Cooper, R. (2002). *Modeling High-Level Cognitive Processes*. Lawrence-Erlbaum Assoc.

Cooper, R. & Shallice, T. (1995). SOAR and the case for unified theories of cognition. *Cognition, 55 (2)*, 115-149.

Duhem (1954). *The Aim and Structure of Physical Theory*. Princeton: Princeton University Press.

Elman (1998). Connectionism, artificial life, and dynamical systems. In W. Bechtel & G. Graham (Eds.), *A Companion to Cognitive Science* (pp. 502-503). Massachusetts: Blackwell Publishers Inc.

Gelman, R. (1982). Accessing one-to-one correspondence: Still another paper about conservation. *British Journal of Psychology, 73*, 209-220.

Giere, R. (1988). *Explaining science: A Cognitive approach*. Chicago and London: The University of Chicago Press.

Hughes, R.I.G. (1997). Models and Representation. *Philosophy of Science, 64*, 325-336.

Hempel, C.G. (1965). *Aspects of scientific explanation, and other essays in the philosophy of science*. Free Press: New York.

Kieras, D. E. (2004). *EPIC Architecture: Principles of Operation*. Retrieved August 28, 2007 from: ftp://www.eecs.umich.edu/people/kieras/EPICtutorial/EPICPrinOp.pdf

Laird, J., Newell, A. & Rosenbloom, P. (1987). SOAR: An Architecture for general intelligence. *Artificial Intelligence*, *33*, 1-64.

Lehman, J.F., Laird J.E. and Rosenbloom, P. (1996). A Gentle introduction to Soar: An Architecture for human cognition. In Scarborough and Sternberg (Ed). *An Invitation to cognitive science: Methods, models and conceptual issues* (211-253). London: MIT Press.

Lehman, J.F., Laird J.E. & Rosenbloom, P. (2006). *A Gentle introduction to Soar: An Architecture for human cognition*. Retrieved March 15, 2007 from: http://ai.eecs.umich.edu/soar/sitemaker/docs/misc/GentleIntroduction-2006.pdf

McCloskey, M. (1991). Networks and theories: The place of connectionism in cognitive science. *Psychological Science, 2 (6),* 387-395.

Morgan, M.S. & Morrison, M. (1999). *Models as mediators* (Ed). Cambridge: Cambridge University Press.

Nagel, E. (1979). *The structure of science; problems in the logic of scientific explanation*. Indianapolis.

Nason, S. & Laird, J. (2004). Soar-RL: Integrating reinforcement learning with Soar. *Cognitive Systems Research, 6*, 51-59.

Newell, A., Simon, H. (1976). Computer science as empirical enquiry: Symbols and search. *Communications of the ACM, 19(3)*, 113-126.

Newell, A. (1990). *Unified theories of cognition*. Cambridge, MA: Harvard University Press.

Piaget (1964). Development and learning. In R.E. Ripple & V.N. Rockastle (Eds), *Piaget rediscovered*. Ithaca, NY: Cornell University Press.

Polk, T.A., & Seifert, C.M. (Eds). (2002). *Cognitive Modeling*. London: The MIT Press.

Rosenbloom, P.S., Laird J.E., Newell, A. and McCarl, R. (1991). A Preliminary analysis of the Soar architecture as a basis for general intelligence. *Artificial Intelligence*, *47*, 289-325.

Rosenbloom, P.S., Laird, J.E., & Newell, A. (Eds). (1992). *The SOAR papers: Research on integrated intelligence*. Cambridge, MA: MIT Press.

Rumelhart, D. E. & McClelland, J. L (1986). *Parallel distributed processing: Explorations in the microstructure of cognition. Vol. 1: Foundations*. Cambridge, MA: MIT Press.

Simon, T., Newell, A. & Klahr, D. (1991). A Computational Account of Children's Learning about Number Conservation. In P.S. Rosenbloom, J.E. Laird & A. Newell (Eds), *The Soar Papers* (pp. 1360-1400). Cambridge: MIT Press.

Suppes, P. (1962). Models of data. In E. Nagel, P. Suppes, A. Tarski (Ed). *Logic, Methodology and Philosophy of Science, Proceedings of the 1960 International Congress* (252-261). Stanford.

Suppe, F. (1977). *The Structure of Scientific Theories*. Urbana: University of Illinois Press.

Suppe, F. (1989). *The Semantic Conception of Theories and Scientific Realism*. Urbana and Chicago: University of Illinois Press.

Suppes (1960). A Comparison of the Meaning and use of models in mathematics and in the empirical sciences. *Synthese, 12*, 287-301.

Taatgen, N. (1999). *Learning without limits: From problem solving towards a unified theory of learning*. Retrieved August 28, 2007 from http://www.ai.rug.nl/~niels/thesis/index.html.

Van Fraasen, B. (1980). *The Scientific Image*. Oxford: Clarendon Press.