PLAUSIBLE REASONING: AN OUTLINE OF THEORY AND EXPERIMENTS

by

R. S. Michalski
K. Dontas
D. Boehm-Davis

Chapter in the book, Methodologies for Intelligent Systems, 4, Z. W. Ras (Ed.), North Holland, 1990.

Methodologies for Intelligent Systems, 4

Proceedings of the Fourth International Symposium on Methodologies for Intelligent Systems held October 12–14, 1989, in Charlotte, North Carolina

Editor

Zbigniew W. Ras, Ph.D.
Department of Computer Science
University of North Carolina at Charlotte
Charlotte, North Carolina

North-Holland
New York • Amsterdam • London

Elsevier Science Publishing Co., Inc.
655 Avenue of the Americas
New York, New York 10010

Sole distributors outside the United States and Canada:
Elsevier Science Publishers B.V.
P.O. Box 211, 1000 AE Amsterdam, the Netherlands

This book is printed on acid-free paper.

Library of Congress Cataloging-in-Publication Data

International Symposium on Methodologies for Intelligent
Systems (4th : 1989 : Charlotte, N.C.)
Proceedings of the fourth International Symposium
on Methodologies for Intelligent Systems, held
October 12–14, 1989, in Charlotte, N.C.
(Methodologies for intelligent systems ; 4)
"The symposium was...sponsored by UNC-Charlotte...
[et al.]"—Fwd.
1. Artificial intelligence—Congresses. 2. Reasoning—Congresses. 3. Knowledge, Theory of—Congresses.
4. Logic, Symbolic and mathematical—Congresses.
I. Ras, Zbigniew. II. University of North Carolina at
Charlotte. III. Title. IV. Series: Methodologies for
intelligent systems ; v. 4.
Q334.I576 1989 006.3 89-22991
ISBN 0-444-01516-7

Current printing (last digit):
10 9 8 7 6 5 4 3 2 1

Manufactured in the United States of America

PLAUSIBLE REASONING: An Outline of Theory and Experiments

(An invited talk)

R. S. Michalski and K. Dontas
Center for Artificial Intelligence

D. Boehm-Davis
Department of Psychology

George Mason University
Fairfax, VA 22030

Abstract

This chapter presents a brief review of a computational theory of human plausible reasoning developed by Collins and Michalski, and discusses experiments conducted toward its validation. This is a descriptive theory that attempts to describe how people actually reason from imperfect premises, in contrast to well-studied normative theories, such as probabilistic reasoning, non-monotonic reasoning, fuzzy logic and multiple-valued logic. The theory proposes a variety of inference patterns that do not occur in formal logic-based theories. It combines semantic and parametric aspects of reasoning, and demonstrates that a large part of human plausible reasoning can be described as small "perturbations" of believed knowledge structures. Ideas are illustrated by the analysis of two protocols recording the explanations of the reasoning process given by human subjects. Preliminary conclusions and directions for future research are presented.

1. Introduction

Unlike in formal logic, premises for reasoning in real-life situations are typically incomplete, uncertain, imprecise or indirectly relevant. Yet, humans have a remarkable ability to reason and derive useful conclusions from such imperfect premises. For example, people can find a desired place in a newly visited city from a combination of sketchy directions from a passer-by, imprecise information in a map, and general knowledge of city organization. They are able to integrate various bits and pieces of information from different sources, resolve contradictions if they occur, and derive the most likely conclusion.

Collins and Michalski (1989) developed a core theory of plausible reasoning that provides a formal framework, a language and a computational model for describing human plausible reasoning processes. It is a descriptive theory that tries to characterize observable aspects of human reasoning, in contrast to normative theories, which treat reasoning as a formal mathematical theory (e.g., Smets et al., 1989). The normative theories are strongly anchored in formal logic, and include probabilistic reasoning (Pearl, 1988; Nilsson, 1986), nonmonotonic reasoning (McCarthy, 1980), default reasoning (Reiter, 1980), fuzzy logic (Zadeh, 1965), and multiple-valued logic (Lukasiewicz, 1967). The primary objective of these theories is to investigate parametric aspects of reasoning, i.e., to develop methods for determining the certainty of conclusions on the basis of the certainty of the premises, without investigating the meaning of the premises. In contrast, the proposed theory attempts to investigate semantic aspects of reasoning, and combine them with parametric aspects. The latter are captured by a collection of different parameters that have influence on the certainty of reasoning, such as typicality, frequency, dominance, dependency, etc. The theory includes a variety of inference patterns that do not occur in formal logic-based theories.

Methodologies for Intelligent Systems, 4
Zbigniew W. Ras, Editor

The efforts on the development of a descriptive theory has started earlier by such researchers as Ajdukiewicz (1965) and Polya (1968). This chapter gives a brief overview of the Collins-Michalski theory, and presents some experiments toward its validation. A detailed exposition of the theory is in the report by Collins and Michalski (1989). An early implementation and efforts on various experimental applications are described by Baker, Burstein and Collins (1987), Zemankova and Dontas (1988), Dontas (1988), and Kelly (1988). Further work on the theory and a description of experiments are described by Michalski, Boehm-Davis and Dontas (1989).

2. Components of the Theory

The theory by Collins and Michalski (1989) offers a framework for characterizing recurrent patterns in human reasoning. These patterns have been captured in a model that contains a set of primitives, operators, and basic inference rules that are applied to knowledge residing in a hierarchical representation system. The primitives enable the specification of knowledge components. The operators allow specification of transformations that can be applied to the basic components in the process of plausible inference.

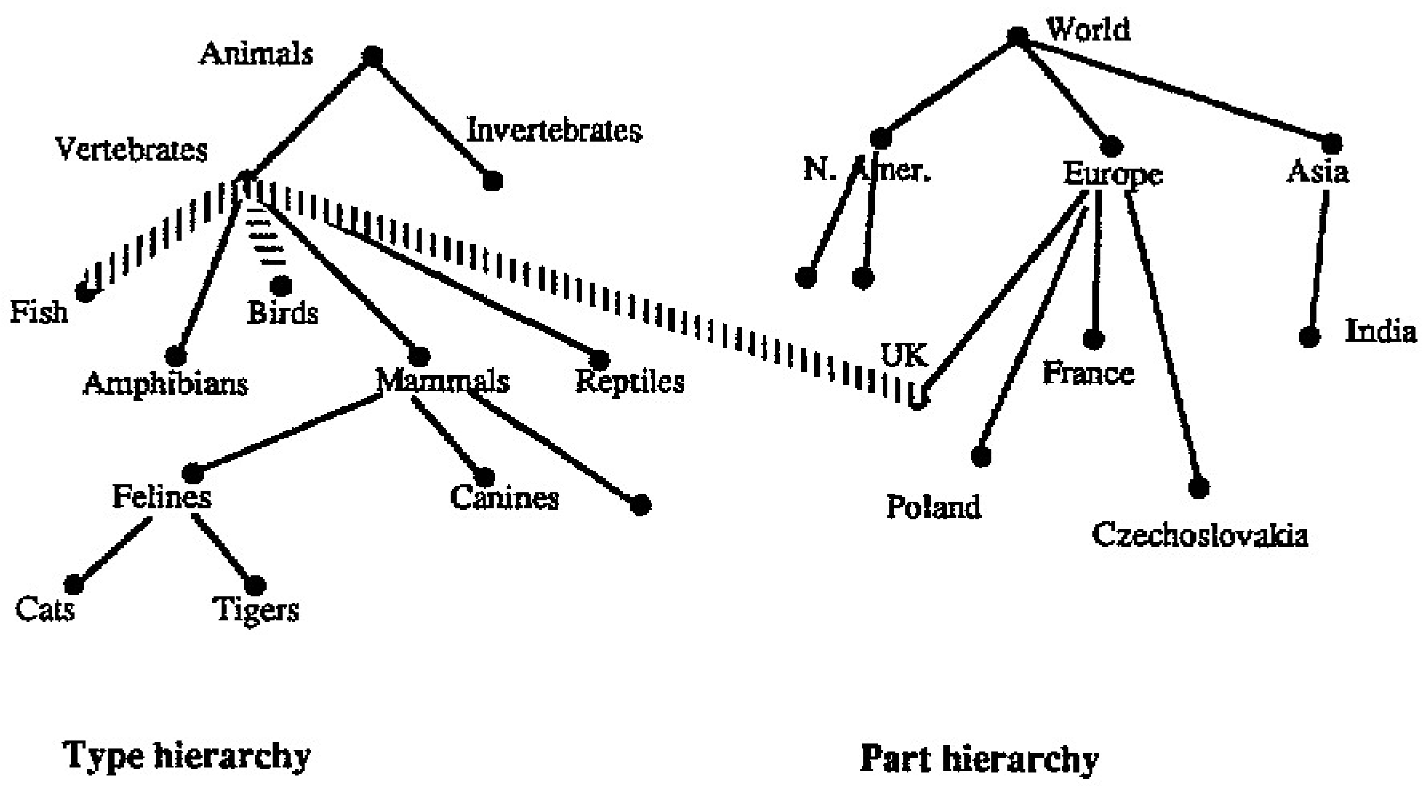

Figure 1: Example Hierarchies and a Trace

Primitives include arguments, descriptors, and referents, which are represented as nodes of *type* (is-a) hierarchy or *part* hierarchy (Figure 1). The hierarchies are dynamic, in the sense that they are growing and changing with experience. Arguments and referents stand for entities (objects, processes, ideas, etc.) in a statement. The same entity may serve as an argument in one statement and as a referent in another. Descriptors are attributes, functions or relations that are used for describing entities. A term is defined as a descriptor applied to one or more arguments, and evaluates to a referent. Simple statements are represented as traces. For example, Figure 1 shows a trace representing a statement that the vertebrates of UK include fish and birds. Following are examples of the elements of the formalism of the Collins-Michalski theory.

Arguments: $a_1, a_2, ..$

Examples: *carnation*
GMU
Cornell

Arguments are represented as nodes of a hierarchy.

Descriptors: d_1, d_2...

Examples:	•	attributes	*color*
	•	functions	*distance*
	•	relations	*greater_than (GT), between*

In the above, "color" is an attribute (a zero- or one-argument descriptor) which is applicable to an entity (carnation), and evaluates to a specific value of color (called referent). "Distance" is a two-argument descriptor. Relations among two or more arguments are other forms of a descriptor.

Terms: $d(a_1)$, $d(a_1, a_2..)$

Examples:	• attr(arg)	*color(carnation)*
	• func(arg_1, arg_2..)	*distance(GMU, Cornell)*
	• rel(arg_1, arg_2...)	*GT(population(VA), population(DC))*

Terms are formed by applying descriptors to one or more arguments, and evaluate to a referent. They have a special significance, because many reasoning tasks can be viewed as evaluating terms. Evaluation of a term may take place by following the trace connecting the descriptor and the argument(s), or by instantiating a general rule (mutual implication or term dependency), or by one or more plausible statement transforms, such as generalizing or similizing, as described below.

Referents: r_1, {r_1, r_2, ..} (values of descriptors; can be descriptors themselves)

Examples: *red*
400_miles
true

The above referents correspond to the three terms exemplified above. The color of a carnation can evaluate to red. The distance between Cornell and GMU is 400 miles. Arguments and referents are distinguished by the position they occupy in statements. Like arguments, referents are entities represented as nodes of some hierarchies. An argument can be any node of a hierarchy, a referent can be any node except for the root node, and a descriptor can be any node except for the leaf node.

Arguments, descriptors, and referents are used in the construction of simple statements, term dependencies and mutual implications. Simple statements are used to represent facts and properties of the objects in the knowledge-base. Mutual implications and term dependencies constitute more complex knowledge, which play the basic role in generating plausible inferences. Term dependencies are related to *determinations* described in the report by Russell and Grosoff (1987). They differ from the determinations in that they can represent bidirectional relationships, and can be specified at different level of abstraction.

Simple statements, term dependencies, and mutual implications are represented as traces linking nodes in different hierarchies. The traces are annotated by a set of parameters (denoted below by π) influencing the strength of the belief in the reasoning process. The parameters represent the frequency of usage, reliability of the source of information, dominance and typicality of a subset within a set, the consistency of the trace with other parts of the knowledge base, the strength of forward and backward implication or term dependency, etc. (Collins and Michalski, 1989).

In this presentation, we will ignore the role of the above parameters, and concentrate primarily on the aspects related to the structural properties of knowledge.

Simple Statements (SS):

$$d(a_1) = r_1 : \pi$$

Examples:

Density(aluminum) = 2.7: π
Age(John) = 55: π
Likes(Robert, Mary) = very_much: π

Term Dependency

$$d_1(a_1) \longleftrightarrow d_2(a_1): \pi$$

Example: Assets(firm) <---> Credit_rating(firm): π

Mutual Implications (MI):

$$SS_i \Longleftrightarrow SS_j: \pi$$

Example: Latitude(place) = north <==> Temp(place) = cold: π

One of the major results of the theory is that plausible inferences correspond to "small perturbations" of the traces. For example, the trace "The vertebrates of UK include fish and birds" (Figure 1) can be used as a base statement for generating inferences "The vertebrates of *Europe* include fish and birds" (a deductive generalization), or that "The vertebrates of Sussex (a part of UK) include fish and birds" (an inductive specialization). Depending on the the direction and size of perturbation, the result of inference may decrease or preserve the certainty. For example, the inductive specialization mentioned above produces a decrease of certainty (Michalski and Zemankova, 1989).

As stated earlier, hierarchies develop and improve with experience. Experts with a lot of experience have more detailed hierarchies than novices. The "small" perturbations of their hierarchies are therefore smaller than "small" perturbations of the novice hierarchies, and thus their plausible inferences are less likely to go wrong. This may be one reason why experts make better guesses than lay people (Matwin, 1989).

3. Statement Transforms

The theory defines eight basic transforms of a simple statement. These transforms are viewed as forms of plausible inference. A transform is done by "perturbing" the argument or referent in a trace spanning one or more hierarchies. As mentioned above, the plausibility of the resulting statement is dependent of the type of perturbation. It also depends on the parameters associated with the base statement. The transforms are classified into two groups. In the first group, transforms modify the argument, whereas in the second group, they modify the referents. The modification is done by generalizing, specializing, similizing, or dissimilizing. For simplicity, the certainty parameters are omitted in the following examples. To describe the transforms we use the following notation:

Generalization of a node "a" in a hierarchy to another node "a' " in context "CTX" is denoted

a' GEN a in CTX(d(a'))

where d(a') denotes descriptors relevant to a' in the given context.

For example, a bird is a generalization (GEN) of chicken in the context (CTX) of birds and their physical features.

Specialization of a node "a" in a hierarchy to another node "a' " in the context "CTX" is denoted

a' SPEC a in CTX(d(a'))

For example, a chicken is a specialization (SPEC) of fowl in the context (CTX) of fowl and their general properties.

The fact that a node "a" in a hierarchy is similar to another node "a' " in the context "CTX" is denoted

a' SIM a in CTX(d(a'))

For example, ducks are similar (SIM) to geese in the context (CTX) of physical features of birds.

The fact that a node "a" in a hierarchy is dissimilar from another node "a' " in the context "CTX" is denoted

a' DIS a in CTX(d(a'))

For example, ducks are dissimilar (DIS) from geese in CTX of neck-length of birds.

Before we formally describe the eight transforms, Figure 2 gives an example of each transform as applied to the base statement: "Flowers of England include daffodils and roses."

BASE STATEMENT: Flower-type(England) = {daffodils, roses, ..}

GEN-A	*(Generalizing Argument)*	Flower-type(Europe)	= {daffodils, roses,..}
SPEC-A	*(Specializing Argument)*	Flower-type(Surrey)	= {daffodils, roses,..}
SIM-A	*(Similizing Argument)*	Flower-type(Holland)	= {daffodils, roses,..}
DIS-A	*(Dissimilizing Argument)*	Flower-type(Brazil)	≠ {daffodils, roses,..}
GEN-R	*(Generalizing Referent)*	Flower-type(England)	= {temperate flowers}
SPEC-R	*(Specializing Referent)*	Flower-type(England)	= {yellow roses}
SIM-R	*(Similizing Referent)*	Flower-type(England)	= {peonies, ..}
DIS-R	*(Dissimilizing Referent)*	Flower-type(England)	≠ {bougainvillea, ..}

Figure 2: Examples of Statement Transforms

A simple statement can be a seed for four different type of inferences: generalizing, specializing, similizing and dissimilizing transforms. Each type can be applied either to an argument or a referent, thus we have a total of eight transforms.

Generalizing Argument (GEN-A)

Generalizing argument extends the applicability of a descriptor-referent pair from an argument to its ancestor. The confidence in the generalized statement is less than in the base statement (Michalski and Zemankova, 1989). The validity of the transform essentially depends on the predictability of the descriptor value from a general node to a specific node, and the typicality of the more specialized argument within the more generalized node, and multiplicity of arguments. The predictability of the descriptor value is proportional to the uniformity of the referent among specialized nodes. In the examples given below, formal ways of using and combining various parameters are not addressed.

$$\frac{\text{Descriptor}(\text{Argument}_1) = \text{Referent} \quad \text{Argument}_2 \text{ GEN Argument}_1 \text{ in CTX} \quad \text{Descriptor} \longleftrightarrow \text{CTX}}{\text{Descriptor}(\text{Argument}_2) = \text{Referent}}$$

Example:

Given:

Performance (Unisys, 1988) = good
Computer_companies GEN Unisys in CTX(Business_type)
Performance <--> Business_type:

Conclude:

Performance(Computer_companies, 1988) = good

In the above example, the base statement says that the performance of Unisys in 1988 was good. Unisys is represented in the hierarchy of companies and the node corresponding to computer_companies is its ancestor. The typicality of Unisys within computer_companies is high. There is also a term dependency which states that business_type of a company is relevant to the performance of a company. Using all this information, we can generalize the base statements to infer that it is likely that the performance of computer_companies in 1988 was good.

Specializing Argument (SPEC-A)

In contrast to the generalizing argument transform, the specializing argument transform restricts the scope of a descriptor-value. If the descriptor-value were to be inherited from a generalized node to the specialized node without exceptions, the inference would be deductive and certain. The statement "*mammals have four legs* " would imply that the *kitty cat* (who is a mammal) *has four legs*. The formalization of the specialization transform goes beyond a mere deductive inference and attempts to look for exceptions by validating the inference after ascertaining that the inheritance of the descriptor value is justified.

For example, in the process of assigning "*four legs*" to a whale, the reasoning process would look at the context of "*habitat*", which has a close functional connection to legs (by means of locomotion). It would see that a whale is not a typical mammal with respect to habitat, and therefore the conclusion that "*a whale* (which is a mammal) *has four legs* " would be blocked. A similar analysis would hold for a *bat* which is a mammal, but is atypical with respect to the means of locomotion and habitat among mammals. Notice that such relations between two or more descriptors can be used in multiple ways.

For example, it can be easily deduced that "*a tiger* , which is a mammal, *has four legs*." However we cannot infer that "*a tiger has claws,*" since the rule that "*mammals have claws*" is too weak. However, such an inference can be strengthened by noting that "a *tiger is a hunting animal.*" Since there is a close functional relationship between claws and hunting activity, one might deduce that "*a tiger has claws*." Note that the same line of reasoning would allow an inference that "*an eagle*, which is a bird of prey, *has claws*," on the same grounds of functional association, though *eagle* and *tiger* are otherwise far removed in the type hierarchy of animals than *tiger* and *cow*.

The strength of the inference depends on the background knowledge as to the alternative means of hunting. There is a need to combine not just one, but several lines of reasoning, as is clear from a parallel example that "the *tigers have sharp teeth*" but "the *eagles have no teeth at all!!*" The further one is away from the base statements, the more one has to look for alternative explanations and new evidence.

$$\frac{\text{Descriptor}(\text{Argument}_1) = \text{Referent} \quad \text{Argument}_2 \text{ SPEC Argument}_1 \text{ in CTX} \quad \text{Descriptor} \leftrightarrow \text{CTX}}{\text{Descriptor}(\text{Argument}_2) = \text{Referent}}$$

Example:

Given:

Major_religion(So_Amer_Cntries) = {Roman_Catholic, ..}
Brazil SPEC So_Amer_Cntries in CTX(Geo_location)
Major_religion <--> Geo_location

Conclude:

Major_religion(Brazil) = {Roman_Catholic,..}

In the above example, we have a base statement that the major religion in South American countries is Roman_Catholicism. Brazil appears as a lower level node (descendant) of South America in the part hierarchy of places. There is a term dependency stating that religion of a country is related to the geographical location of the country (countries in the same geographical proximity tend to have similar religious background). From this it can be concluded that the major religion in Brazil is Roman_Catholicism.

Similizing Argument (SIM-A)

The similizing argument is a statement transform which depends on the similarity between two arguments rather than ancestor-descendant relation between them. Because potentially all the nodes in the hierarchy can be used as similar nodes, all the nodes in the hierarchy would need to be examined in order to find the best match. This makes the transform a computationally unattractive means of answering queries unless a good similar argument is known beforehand. This transform is valuable in verifying inferences by other lines of reasoning.

$$\frac{\begin{array}{l}\text{Descriptor}(\text{Argument}_1) = \text{Referent}\\ \text{Argument}_2\ \text{SIM}\ \text{Argument}_1\ \text{in CTX}\\ \text{Descriptor} \leftarrow\!\!\rightarrow \text{CTX}\end{array}}{\text{Descriptor}(\text{Argument}_2) = \text{Referent}}$$

Example:

Given:

Economic_state(Singapore) = Excellent
Hong Kong SIM Singapore in CTX(Economy_type, Tax, Latitude, Resources, Communication, ..)
Economic_state <--> CTX

Conclude:

Economic_state(Hong Kong) = Excellent

This example uses the similarity between argument to deduce that economic_state of Hong_Kong is strong. The inference is based on the information that economic_state of Singapore is excellent, that Hong_Kong is very similar to Singapore in the feature space of economy_type, tax, resources, communication, and that feature space is relevant to the economic_state of a country.

Dissimilizing Argument (DIS-A)

The dissimilizing argument transform depends on the dissimilarity between two arguments. The transform depends on the assumption that if some context is relevant to the descriptor, then two arguments which are dissimilar in the context will likely have different descriptor-value (referent). This transform can be used to eliminate one or more contending hypotheses. It can also be used to increase certainty of a conclusion by showing that alternative hypotheses are not plausible.

$$\frac{\begin{array}{l}\text{Descriptor(Argument}_1\text{) = Referent}\\ \text{Argument}_2\text{ DIS Argument}_1\text{ in CTX}\\ \text{Descriptor <--> CTX:}\end{array}}{\text{Descriptor(Argument}_2\text{)} \neq \text{Referent}}$$

Example:

Given:

Carnivorous(Tiger) = yes
Tiger DIS Cow in CTX(sharp_teeth, claws, ..)
Carnivorous <--> CTX

Conclude:

Carnivorous(Cow) ≠ yes

This example uses the dissimilarity between argument to deduce that a cow is not a carnivorous animal. The inference is based on the premises that cow and tiger differ with regard to having or not having sharp teeth and claws, and that these properties are important for carnivorous animals.

4. Experimental Study

Studies were conducted with human subjects who were asked to answer questions requiring them to conduct reasoning. Their answers were analyzed in terms of the concepts and inference rules developed in the theory. The purpose of this analysis was to validate the theory and to determine what enhancements or extensions might be needed to account for the data. This analysis was restricted to the structural properties of the model, and the types of inferences involved in reasoning. Future studies will examine the processes associated with assigning certainty to the conclusions.

The Collins-Michalski theory was initially developed using inferences that people made about specific domains with a well specified, small knowledge base where the participants had no special knowledge about the domain within which they made the inferences (e.g., reasoning about weather patterns in a geographical domain; Collins, 1978).

The goal of the current research is to determine whether the theory is adequate for describing a set of inferences generated by subjects in response to a set of questions about a new domain. The experiments were also designed to examine the impact of world knowledge on the inference process. In the current study, people were asked to make inferences about a domain where they may have some prior knowledge that could be brought to bear on the inference process.

A table composed of 13 countries and attributes for characterizing these countries were used in the study. For each country, the descriptors, such as the type of government, the type of press, literacy rate, type of work force, major religions, trading partners, major industry, per capita income, and relations with the United States were determined from published literature. For the purpose of the experiments, the 18 of the country attribute values were replaced with question marks. These attribute values were the characteristics that the subjects were asked to infer in the experiment. Another version of the table was created in which the country names were replaced with three letter nonsense names (e.g., ABC, DEF). Subjects who received the second table were not told that the rows in the table represented actual countries.

The participants were provided with a copy of one of the two versions of the table (four subjects received table with the actual country names, the other four received a table with the nonsense names). The nature of the table was explained to the participants. They were then asked to generate plausible entries for each of the cells which contained a question mark. Thus, they were asked to make a plausible inference for each of 18 cells in the table. They were asked to verbalize their thought processes and the reasons for their conclusions as they completed their task. Verbal protocols were recorded and transcribed for analysis.

5. Selected Protocols and their Analyses

Two examples out of 144 protocols collected are given below. They were selected to represent non-trivial, typical, and interesting reasoning cases. In the protocols, RS means a reasoning step, PBK means personal background knowledge, GBK indicates given background knowledge (table), MI indicates inference from mutual implication, M Recall means memory recall. The present analysis is preliminary and may be improved in further studies.

Question 10B:

What are the major religions in Canada?

Subject

Canada. Um, well, Canada is split between the French speaking sector, as well as English speaking sector, which given those two warring factions and how that conflict rather manifests itself in the language debate. Should the official language be French or should it be English. Um, given how language is so closely tied to religion, I imagine that it's probably Protestant versus Catholic, as well. Although that is not an issue that surfaces so much, that's my thought. So it's probably two religions.

Verbal analysis

The subject notes the existence of warring factions in Canada based on language. The subject also notes that there is close dependence between language and religion. French speaking people tend to be Roman Catholic, and English speaking people tend to be Protestants. For these reasons the subject concludes that the major religions of Canada include both Roman_Catholic and Protestant.

Formal analysis

RS1
Lang(people(Canada)) = {French, English,..}: — M Recall

RS2
Lang(people(Country)) <--> Maj_religion(people(Country)) — PBK

RS3
Lang(people(Canada)) = {French, ...} <==> Maj_religion(people(Canada)) = {R. Cath,..} — PBK
Lang(people(Canada)) = {French,..} — PBK

Maj_religion(people(Canada)) = {R Cath, ..} — MI

RS4
Lang(people(Canada))={ English, } <==> Maj_religion(people(Canada))={Prot,.. } — PBK
Lang(people(Canada)) = {English,..} — PBK

Maj_religion(people(Canada)) = {Prot, ..} — MI

Conclude: Maj_religion(Canada) = {R. Cath, Prot,..},

Question 8B:

What is the type of work force in Vietnam?

Subject

Vietnam. Work force. I think it is primarily agricultural. It is way behind Pacific Rim, the development of the rest of the Pacific Rim countries because of the Vietnam war. And the continuing state of, it is very poor. The refugees, there was a mass exodus of refugees, a brain drain, if you will, during the war, after the war, continuing still. Therefore that does not leave a lot of room to revolutionize, to modernize what little industry you might have, that might have survived the war. Uh, I think it is primarily agricultural.

Verbal analysis

The subject starts with a weak recall that the type of work force in Vietnam is primarily agricultural. Then the subject offers justifications. Development of Vietnam has lagged behind its neighbors in the Pacific Rim area due to war. The economic status of Pacific Rim countries is high and Vietnam's economic situation is worse than theirs. Therefore, the economic status of Vietnam is likely to be poor. If a country has undergone recent war, if its economic status is not sound, if there are refugees from the country, then the development of the country is slow. Such a situation existed in Vietnam, therefore, its development is slow.

If a country is trying to modernize, then it will change its industry from agricultural to modern industry. This change is slow if the development of the country is slow in general, therefore change from agricultural to modern industry is slow in case of Vietnam. If there is no modern industry, then there is no large scale industrial force, therefore, the work force of Vietnam is not industrial but agricultural, therefore, the subject infers that the labor force of Vietnam is predominantly agricultural, which is consistent with the earlier recall.

Formal analysis

RS1

Labor_force(Vietnam) = {agric, ..}	M Recall

RS2

Mil_stat(Cntry)=war <==> Econ_stat(Cntry) < Econ_stat(Nbors(Cntry))	PBK
Mil_status(Vietnam) = war	PBK
Nbors(Vietnam) = Pacific_Rim_cntries	PBK
Econ_stat(Vietnam) < Econ_status(Pacific_rim_cntries)	MI

RS3

Econ_stat(Pacific_Rim_cntries) = high	PBK-Implicit
Econ_stat(Vietnam) < Econ_stat(Pacific_Rim_cntries)	RS2
Econ_stat(Vietnam) = not high	MI

RS4

Mil_stat(Cntry) = at_war <==> Xodus(Cntry) = high & Brain_drain(Cntry) = high	PBK
Mil_stat(Vietnam) = at_war	GBK
Xodus(Vietnam) = high & Brain_drain(Vietnam)= high	MI

```
RS5
Xodus(Cntry) = high & Brain_drain(Cntry)= high <==> Develop(Cntry) = slow      PBK
Xodus(Vietnam) = high & Brain_drain(Vietnam)= high                             RS4
------------------------------------------------------------
Develop(Vietnam) = slow                                                        MI

RS6
Tendency(Cntry) = modernize <==> Changeto(Agric, mod-ind)                      PBK
Develop(Cntry) = slow <==> Change(agric, mod-ind) = slow                       PBK
Develop(Vietnam) = slow                                                        RS5
------------------------------------------------------------
Changeto(agric, mod-ind) = slow

RS7
Change(agric, mod-ind) = slow                                                  RS6
Industry(Vietnam) = {agric, ..}                                                GBK
Industry(Cntry) = {agric, ..} <==> Lab_force(Cntry) = {agric, ..}              GPK
------------------------------------------------------------
Lab_force(Vietnam) = {agric, ..}                                               MI

Conclude     Lab_force(Vietnam) = {agric, ..}
```

In collecting the protocols, the subjects were briefly told the purpose of the experiment. No specific time limit was set to answer the questions. The subjects typically took between an hour to answer the 18 questions. These set of protocols generally emphasized simple reasoning patterns involving reasoning by application of one or more mutual implications. Reasoning patterns involving constructive processes, such as discovery of dependencies or checking for consistency of personal knowledge with that available in the table were absent.

6. Conclusions and Open Questions

Preliminary analysis confirmed that people follow several lines of reasoning in reaching a conclusion. The individual lines are weighted and compared. If different lines lead to different conclusions with a similar weight, a subject does not express any opinion ("I do not know"). The study has also showed that some needed rules were not captured in the original model. A detailed report on the experiments conducted is in (Michalski, Boehm-Davis and Dontas, 1989).

Hierarchies, term dependencies and mutual implications are very important components of the process of plausible reasoning. In the present study, the question of how people learn these components was not addressed. Further research needs to be done to find a computational model of how people create conceptual hierarchies, and discover implications and dependencies. The theory also needs to be related to existing methodologies, and extended to include temporal reasoning, spatial reasoning, reasoning under time and resource constraints (e.g., related to the variable precision logic, as described by Winston and Michalski, 1986), as well as meta-knowledge reasoning.

It was seen in the protocols that the subjects always attempted to build a consistent, plausible scenario to explain their conclusions based on beliefs and personal background knowledge (PBK). The PBK was occasionally false, or invented for the purpose of answering the question. Subjects that know country name rely primarily on PBK, rather than on GBK, i.e., on the given background knowledge (the table).

In the current study, the influence of various parameters (certainty, typicality, frequency, etc.) on the reasoning process was not taken into account. This will be done in the next stage of research. APPLAUSE (Dontas and Zemankova, 1988) and PRS (Kelly, 1988), two systems implementing some aspects of Collins-Michalski theory will be modified to perform similar tasks, and the results, conclusions and lines of reasoning demonstrated by people will be compared with those used by the system.

In conclusion, the experiments have demonstrated that the theory provides an adequate mechanism for representing reasoning for the class of tasks investigated. The theory offers new tools for knowledge representation, and has a potential for applications in a variety of fields, such as decision making and analysis, diagnosis (medical, agricultural or technical), goal recognition, intelligent tutoring, object and scene recognition, planning, autonomous robotics, estimating costs and labor in design, document retrieval systems, etc.

Acknowledgements

The authors thank Dr. Allan Collins, Professor Doug Medin, and Dr. Maria Zemankova for their collaboration and contributions at various stages of the project, and Joyce Ralston for valuable help and proofreading of the article. This research was done in the Artificial Intelligence Center of George Mason University. Research activities of the Center are supported in part by the Office of Naval Research under grant No. N00014-88-K-0397, in part by the Office of Naval Research under grant No. N00014-88-K-0226, and in part by the Defense Advanced Research Projects Agency under grant, administered by the Office of Naval Research, No. N00014-87-K-0874.

REFERENCES

Ajdukiewicz, K., *Logika Pragmatyczna*, Panstwowe Wydawnictwo Naukowe, 1965.

Baker, M., Burstein, M. H., and Collins, A. (1987), Implementing a Model of Human Plausible Reasoning, *Proceedings of the Tenth International Joint Conference of Artificial Intelligence*. Vol. 1 (pp. 185-188), Morgan Kaufman, Los Altos, CA.

Collins, A. (1978). Human Plausible Reasoning, Report No. 3810, Bolt Beranek and Newman Inc., Cambridge, MA.

Collins, A. and Michalski, R. S. (1989), The Logic of Plausible Reasoning: A Core Theory, *Cognitive Science*, Jan 1989, pp. 1-49.

Dontas, K. & Zemankova, M. (1988), APPLAUSE: An Implementation of the Collins-Michalski Theory of Plausible Reasoning, *Proceedings of the Third International Symposium on Methodologies for Intelligent Systems*, Torino, Italy, 1988.

Kelly, J. (1988), PRS: A System for Plausible Reasoning, M.S. Thesis, Department of Computer Science, University of Illinois, Urbana.

Lukasiewicz, J. (1967), Many-valued Systems of Propositional Logic, In *Polish Logic*, S. McCall (Ed.), Oxford University Press.

Matwin, S. (1989), Personal communication.

McCarthy, J. (1980), Circumscription - A Form of Non-monotonic Reasoning, *Artificial Intelligence*, 13 (1,2), pp. 27-39.

Michalski, R. S. & Winston. P. H. (1986), Variable Precision Logic, *Artificial Intelligence Journal*, 29, 121-146.

Michalski, R. S., Boehm-Davis, D. and Dontas, K., (1989), Theory of Human Plausible Reasoning: Fundamental Ideas and Experiments, to appear in *Reports of Machine Learning and Inference Laboratory*, AI Center, George Mason University, 1989.

Michalski, R. S. and Zemankova, M. (1989), to appear in *Reports of Machine Learning and Inference Laboratory*, AI Center, George Mason University, 1989.

Nilsson, N. J. (1986), Probabilistic Logic, *Artificial Intelligence*, 28, 71-87.

Pearl, J. (1988), *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, Morgan Kaufmann.

Polya, G. (1968), *Patterns of Plausible Inference*, Princeton University Press.

Reiter, R. (1980), Logic of Default Reasoning, *Artificial Intelligence*, 13, 1-132.

Russell, S. J., and Grosoff, B. N. (1987), A Declarative Approach to Bias in Concept Learning, *Proceedings of AAAI-1987*, pp. 505-10, Morgan Kaufmann.

Smets, P., Mamdani, A., Dubois, D. and Prade, H. (1988), *Nonstandard Logics for Automated Programming*, Academic Press.

Zadeh, L. A. (1965), Fuzzy Sets, *Information and Control*, 8, pp. 338-353.