ISG 1978-6

1978-6

78-6

A DISCUSSION OF PLAUSIBLE INFERENCES AND A FORMALIZATION OF SOME PLAUSIBLE INFERENCE RULES

Mitchell Lubars

December 1978

Internal Report

Throughout life, people are constantly faced with the task of making decisions. As infants, we decide when we must attract the attention of our parents to attend to our physiological needs, as well as our emotional needs. In later years, we learn to make more complicated decisions. That is, we develope the ability to take more variables into account, which will affect the outcome of our decisions.

Some of the decisions which we make are clear cut; there is no doubt in our minds that the evidence we observe implies that the decision must be made with a particular outcome. The infant who falls out of his crib, has no doubt that he must attract the attention of a parent who will see to his needs. And the mother who hears her baby crying has no doubt that the baby needs her attention for one reason or another.

Sometimes the decisions which we make do not involve actions, but are decisions that involve observations. We decide what has happened, is happening, or will happen in the future. Such decisions involve justification, analysis, or prediction and may also be clear cut decisions. We predict that if we hit our hand with a hammer, we will sense pain. There is no doubt associated with our prediction, for we know that pain is an inevitable consequence of such an action.

We consider all such decisions to be inferences, and we infer that consequences result when a conjecture is satisfied. If the baby is in pain, a consequence will be that he attracts his parent's attention; He cries. The mother hears the baby crying, and consequently she will see to its needs. In making inferences, like the ones discussed, we apply inference rules. So far, all of our examples have been applications of a particular inference rule. Stated briefly, the rule is, If A implies B and A is true, then infer B. The mother applies the rule; the baby cries because it needs my attention. Since it is now crying, it must need my attention.

Life would be full of simple decisions if all inference rules were as cut and dry as the above rule. We all know from experience that this just isn't true. In fact most decisions, can not be made with absolute certainty at all; some decisions we make are just plain wrong. How do we cope with situations in which an inference must be made with less than absolute certainty? What kind of such inferences can we make?

We begin our discussion with a few definitions. Inferences which can always be made with absolute certainty are called "Demonstrative Inferences". "Heuristic Inferences" are inferences which can only be made with less than absolute certainty. "Plausible Inferences" involve both demonstrative and heuristic inferences. Their applications result in certainties within the continuous range from no certainty to absolute certainty, the boundaries included. The baby knows that if he is hungry, he should cry to attract his parent's attention. He thinks that he is a little bit hungry, and so he thinks that he might want to cry. The mother knows that if the baby is crying, the baby needs her attention. She thinks that the baby is crying, but she is not sure. The baby may need her attention.

Application of one or more inferences in order to reach a decision is called, reasoning. Plausible reasoning involves the application of one or more plausible inferences. In demonstrative reasoning, it is sufficient to apply a single chain of inferences in order to infer the desired consequence with absolute certainty, assuming that such a series can be found and is applicable at each successive step. Such a chain constitutes a proof. In plausible reasoning, (more correctly, heuristic reasoning), a single such chain of inferences is only enough to gain some insight into the consequence. In order to get a better idea of what certainty can be assigned to the consequence, it is necessary to follow as many such plausible inference chains as possible. Such, investigation may lead to a consequence with much higher certainty, although contradictory consequences may just as easily be inferred.

This paper will examine plausible inferences in two lights. The work of George Polya[1] in the application of plausible inferences to the demonstration of mathmatical conjectures will be discussed and some of his rules will be summerized. Then the work of Allan Collins[2, 3] in his investigation of human plausible reasoning will be discussed, and some of his rules will be presented followed by a more mathematical formalization of them. In neither case, will the mechanism of plausible reasoning be discussed, beyond the application of individual plausible inference rules. For such examples of plausible reasoning, see [3].

Frequently, a proof of a mathematical conjecture is either difficult or impossible to find. It may be, however, that we can gain some insight into the problem by observing that conjecture, A, implies a consequence, B. Or perhaps A is a consequence of another conjecture, B. Or it may be that a conjecture B is found to either be analogous to A or to be a rival conjecture. In each of these instances, while we are unable to determine anything about A, we still may be able to determine the truth of B. What kind of credibility can we attribute to A, in either proving or disproving B? What if we can not prove B either, but we can gain some insight into its credibility?

Polya [1] discusses at length various inferences which may be made with absolute certainty, and inferences which may be made with less than absolute certainty in the province of mathematical conjectures. He further discusses how varying beliefs in alternate, and similar conjectures and consequences affect the credibility of an original conjecture. I now present a brief summary of some of these plausible inferences.

Let us say that we have some clearly formulated conjecture, A, and a consequence of A, B. We do not know whether either A or B is true, but we do know however that :

A implies B.

We decide that determining the truth of B is easier than determining the truth of A. What do we stand to learn about A, by determining the truth of B? If B turns out to be fase, we could conclude that A is also false. This pattern of reasoning is the MODUS TOLLENS of the so-called hypothetical syllogism:

```
A implies B
  B false
-----------
  A false
```

The horizontal line separating the two premises from the conclusion stands for the word "therefore". There is no uncertainty about the result in this case. This is a DEMONSTRATIVE INFERENCE.

If B turns out to be true, we can reach no such demonstrative conclusion. Verifying B does not prove A. It does however, make the conjecture A seem more credible, the implication has been verified in the one case. Therefore this is a pattern of PLAUSIBLE inference:

```
A implies B
   B true
-----------
A more credible
```

Polya calls this the FUNDAMENTAL INDUCTIVE PATTERN (or shorter, the INDUCTIVE PATTERN). What this essentially says is that the verification of a consequence renders a conjecture more credible.

Suppose that we have already verified the given implication for the consequences B_1, B_2, ... B_n. Now we are able to prove the consequence B_{n+1}, where B_{n+1} is very different from the formerly verified consequences. Since conjecture A had withstood a more extreme test than the verification of another similar consequence, we might consider A to be much more credible. Alternatively, had consequence B_{n+1} been very similar to the previously proven consequences, we might think that its verification would add very little to our confidence in A. We state these new patterns as follows:

A implies B B_{n+1} is very different from the formerly verifed consequences $B_1, B_2, \ldots, B_n$ of A	A implies B B_{n+1} is very similar to the formerly verified consequences $B_1, B_2, \ldots, B_n$ of A
A much more credible	A just a little more credible

The verification of a new consequence counts more or less according as the new consequence differs more or less from the formerly verified consequences.

If it turns out that B was almost a forgone conclusion without the consideration of A, then the verification of B cannot make A seem very much more credible. On the other hand, if B is highly unlikely without the consideration of A, then the verification of B as a consequence of A renders the conjecture much more likely. We add the following patterns to our plausible inferences:

A implies B B very improbably in itself	A implies B B quite probably in itself
A very much more credible	A just a little more credible

The verification of a consequence counts more or less according as the consequence is more or less improbable in itself.

If instead of being able to prove or disprove B, we are only able to either increase of decrease our confidence in B, that is we find B either more or less credible, we can still make some inference about A. The following are the two such patterns:

A implies B B less credible	A implies B B more credible
A less credible	A somewhat more credible

Our confidence in a conjecture is influenced by our confidence in one of its consequences and varies in the same direction.

Let us now say we are given two analogous conjectures A and B. We wish to verify conjecture A if possible, but we suspect that we can more easily prove conjecture B. What insight into conjecture A can be found if indeed conjecture B turns out to be true? We expect the following pattern of plausible reasoning:

```
A analogous to B
   B true
----------------
A more credible
```

A conjecture becomes more credible when an analogous conjecture turns out to be true.

Suppose instead of proving that B is true, that we are only able to increase our belief in the truth of B; we find B to be more credible. We use the following pattern to find A more plausible as well:

```
A analogous to B
B more credible
----------------
A somewhat more credible
```

This is the shaded form of the previous pattern. It states that a conjecture becomes somewhat more credible when an analogous conjecture becomes more credible.

We now consider the converse of our earlier implication. That is what inferences can we make if we know that:

A is implied by B?

If we could prove B, the consequence A would follow. If, however, we succeed in disproving B, A could still be true. There is no demonstrative conclusion. But, we now have one less chance of proving A. A ground for A has had to be discarded, and if there is any change in our confidence in A as a consequence of the disproof of B, it can only be a change for the worse. We observe the following two patterns:

```
A implied by B
   B true
----------------
   A true
```

```
A implied by B
   B false
----------------
A less credible
```

The demonstrative inference is the MODUS PONENS of the so-called hypothetical syllogism. The non-demonstrative, or heuristic, inference can be stated as: our confidence in a conjecture can only diminish when a possible ground for the conjecture is exploded.

Sometimes it is the case that two alternative conjectures A and B are considered, and it is impossible for both of them to be true. We say A conflicts with B or:

A is incompatible with B.

The truth of one conjecture implies the falsehood of the other.

If it turns out that one of the conjucture is proven to be the correct one, let us say B, then the other one must be false. If B is only shown to be more credible, conjecture A is still possible, although it is less plausible. We have the following two patterns:

```
A incompatible with B          A incompatible with B
     B true                          B false
---------------------          ---------------------
     A false                     A more credible
```

Our confidence in a conjucture can only increase when an incompatible rival conjecture is exploded.

When B is found to be less credible instead of false, we have the following shaded pattern of plausible inference:

```
 A incompatible with B
     B less credible
----------------------
A somewhat more credible
```

All these patterns of plausible reasoning, as well as others, may be chained and nested to create new useful patterns. For that reason, the patterns just discussed are considered to be some of the fundamental patterns, by Polya.

Allan Collins investigated the use of human plausible reasoning in attempting to understand the types of protocols people use in answering questions [2]. His analysis attempts to account for such protocols in terms of a taxonomy of plausible inference types, a taxonomy of default assumptions, and what the subject must have know a priori. What Collins' discovered as a result of his questioning and analyzing the protocols of his subjects is the following:

1) There are usually several different inference types used to answer any question.
2) The same inference types recur in many different answers.
3) People weigh all the evidence they find that bears on the question.
4) People are more or less certain depending on the certainty of the information, the certainty of the inferences, and on whether different inferences lead to the same or

opposite conclusions.

The rules described by Collins are plausible inferences which involve mappings from superordinate sets, similar sets, or subordinate sets. Properties may be mapped from one set or from many sets. The different kind of mappings delineated in his theory are:

DEDUCTION (Superordinate Inferences) which map properties of the set onto subsets.
ANALOGY (Similarity Inferences) which map properties from one set to a similar set.
INDUCTION which maps properties of subsets of a set onto other subsets.
GENERALIZATION (Proof-by-cases) which maps properties of subsets of a set onto the set.
ABDUCTION which maps a subset with the same property as some set into the set.

Some of Collins' rules are described in English in [2, 3]. The remainder of this paper deals with developing a formal description of some of the plausible inference rules presented by Collins. The rules dealing with temporal and spatial inferences are intentionally omitted from this discussion since the temporal and spatial properties themselves rely upon other properties which are not pertinent to the other rules, such as transitivity and associativity. In addition, the temporal and spatial rules, as presented by Collins, make various assumptions about the type of such representations required for such properties. This is something that I wish to avoid in developing a formalization at a sufficiently general level.

What follows is a definition of the notation used in the formalization of Collin's rules. The remainder of the paper, after that, is a statement of each of Collins' rules, followed by an example by Collins and conditions that increase the certainty of the rule. For each rule, I then present a formalization of the rule, a formalization of the conditions that increase certainty, and an application of my formalization to Collins' example.

NOTATION USED IN THE FOLLOWING DISCUSSION OF INFERENCE RULES:
-------- ---- -- --- --------- ---------- -- --------- -----

C_i is some concept under consideration, it is the set of all cases of the concept for which it stands for. Note that a case of the concept may be a concept as well.

C^S is the superset, containing all concepts, C, as its elements. Therefore a collection of concepts is a subset of C^S. A single concept C is an element of C^S. Note that C is also a subset of C^S since, the cases of C are also concepts which are elements of C^S. C will be used in both contexts in the following examples, although care will be taken to make the context clear.

C_i is a concept of the concepts which are being considered. C_1 and C_2 are therefore two such concepts.

C' and C'' are complementary concepts. That is C' and C'' differ is some particular way that is of interest in the example.

c is a subset of concept C. It contains particular cases, c_i, as its elements, which will be defined below.

c' and c'' are complementary subsets of C. That is, they are a partition of C based on some property or properties of interest.

c_i is a particular case of C. It is an atomic concept. That is it is an n-tuple, a property value for each know property of c.

P is a set of properties.

P(C) represents the set of properties of a concept, C, based on the above consideration.

p is a subset of P. It contains the properties of interest in a given example.

p(C) represents a subset of the properties of C, those properties of C that we are interested in.

P' and P'' are complementary subsets of P. That is, they are a partition of P based on some criterion.

p_K represents one of the properties that we are interested in.

$p_K(C)$ represents one of the properties of C that we are interested in.

$p_I(C)$ represents the set of independent variables of concept C.

$p_{i_1}(C),\ldots,p_{i_n}(C)$ represents the elements of $p_i(C)$.

$p_d(C)$ represents the dependent variable of concept C.

P_i is the set of p_i for p_d such that $p_i \in P_i$. That is the set of functional dependencies that lead to $[p_d(C) = r_d]$. Note, although $p_i \in P_i$, $p_i \subset P$.

R represents the ranges of the properties P, that we are interested in. It represents the sets of values that P(C) may attain. Furthermore any or all of the values in the set may themselves be sets. For example [kind(animal) ={insect, mammal, reptile}, [kind(mammal) = {human, ox, horse, ...}], [kind(horse) = ...

r is a subset of R. It represents the set of values that p(c) may attain.

r_i is a subset of R. It represents the set of values that $p_i(c)$ may attain.

v_i is a single element of r_i. At times when r_i in fact has only one element, v_i may be used in place of r_i to express this.

t is a transformation from one rule to another rule. It takes as its argument the original rule and generates an inference rule.

:::> is the rule generation symbol. It separates t and its argument from the generated inference rule, specifying that the rule on the right side of the arrow is generated by t acting on the argument rule.

===> is the functional dependency symbol. It separates the dependent variable (p_d), on the right hand side, from the independent variables (p_is) on the left hand side.

---> is the inference symbol. It means the the right hand side of the arrow may be inferred if the left hand side of the arrow evaluates to true.

U C represents set union of all such sets C, as defined in context.

|s| represents the cardinality of set s.

wrt is an abrieviation for "with respect to".

[*] stands for zero of more independent variables in a functional dependency whose values are either not known, not in the range, or not of interest.

With this backgrounc, we may now define what we mean by a concept more precisely. A concept, C, is equivalent to $r_1 \times r_2 \times \ldots \times r_n$, where $[p_k(C) = r_k]$. That is any given concept is defined in terms of its properties. The cases of C, c_i, are themselves concepts which are non-atomic if $r_1 \times r_2 \ldots \times r_n$ has more than one element. Atomic concepts are called particular cases. It follows that if any r_k for some property, $p_k(c_i)$, is a set, then c_i is a non-atomic concept.

INDUCTION FROM CASES

If a property is in a given range for a number of instances of a set, then infer that the property is in that range for any other instances of the set.

EXAMPLE

Several official spokesmen from country X make public statements more belligerent than their know private views. Thus a recent public statement by an official of country X is likely to be overly belligerent.

CONDITIONS THAT INCREASE CERTAINTY

1) The more instances for which the property is in the given range.
2) The less the variability in the values of the property for the different instances.
3) The more typical the instances for which the property is in the given range.
4) The more similar are the instances over which the property is mapped.
5) The more typical the instances onto which the property is mapped.
6) The more typical the instances which are most certain.

FORMULATION OF RULE

$t(\forall c_i \in c \subseteq C \quad [p_K(c_i) = r_K])$
$\quad ::::> \forall c_i \in c \subseteq C \quad [p_K(c_i) = r_K] \dashrightarrow [p_K(C) = r_K]$

Here each c_i is a particular case of a more general concept C. The transformation produces a rule for property p_K allowing induction from some particular cases, of a more general concept, to that concept.

FORMULATION OF CERTAINTY

1) $\forall c_j \in c' \subseteq C \quad [p_K(c_j) = r_K]$, $\forall c_j \in c'' \subseteq C \quad [p_K(c_j)$ .NOT.= $r_K]$
   $|c'|$ larger wrt $|c''|$ ===> increase in certainty
2) $\forall c_j \in c \subseteq C \quad [p_K(c_j) = v_{Kj}]$, $v_{Kj} \in r_{Kj} \subseteq r_K$

$|r_{Kj}|$ smaller wrt $|r_K|$ ===> increase in certainty

3)

4) $\forall c_j \in c \subseteq C \quad \forall p_\ell \in p' \subseteq P \quad [p_\ell(c_j) = r_\ell]$,

$\forall c_j \in c' \subseteq C \quad \forall p_\ell \in p'' \subseteq P \quad [p_\ell(c_j)$ .NOT.= $r_\ell]$

$|p'|$ larger wrt $|p''|$ ===> increase in certainty

5)

6)

APPLICATION

Given that:

$\forall$ spokesman$_i$ $\in$ (spokesmen with beligerent views) $\subseteq$ (all spokesmen)
[public statement(spokesman$_i$) = beligerent]

We can generate the inference rule:

$\forall$ spokesman$_i$ $\in$ (spokesmen with beligerent views) $\subseteq$ (all spokesmen)
[public statement(spokesman) = beligerent]
---> [public statement(all spokesmen) = beligerent]

which we then apply to infer:

[public statement(all spokesmen) = beligerent]

ABDUCTION

If a property is in a given range for a set, and if the property is in that range for a particular instance, then infer that this instance is a member of the set.

EXAMPLE

Countries with nuclear weapons tend to classify their nuclear physics research and import quantities of uranium or plutonium. Country X is doing both these things, so it may have nuclear weapons.

CONDITIONS THAT INCREASE CERTAINTY
---------- ---- -------- ---------

1) The fewer alternative sets for which the property is in the same range.
2) The more properties of the instance that are in the given range for the set.
3) The less likely that the property is in the given range a priori for any instance.
4) The more likely the property is in the given range for instances of the set.
5) The less likely are the other sets for which the property is in the same range.
6) The less correlated are the properties that are in the given ranges for the instance.
7) The more correlated are the other sets for which the property is in the same range.
8) The more likely is the given set (i.e. the more frequently are instances of the set).

FORMULATION OF RULE
----------- -- ----

$t([p_K(C) = r_K])$ ::::> $[p_K(C) = r_K] \wedge [p_K(c_i) = r_K]$ ---> $c_i \in C$

Here c_i is a concept variable of a more general concept, C. The transformation produces a rule for property, p_K, allowing a generalization from a particular case, c_i, to that of a general case, C, by abduction on that property.

FORMULATION OF CERTAINTY

1) $\forall C' \subseteq C^S$ $[p_K(C') = r_K]$, $\forall C'' \subseteq C^S$ $[p_K(C'')$ .NOT.= $r_K]$,
   $|C'|$ smaller wrt $|C''|$ ===> increase in certainty
2) $\forall p_\ell \in P' \subseteq P$ $[p_\ell(c_i) = r_\ell]$, $\forall p_\ell \in P'' \subseteq P$ $[p_\ell(c_i)$ .NOT.$r_\ell]$
   $|P'|$ greater wrt $|P''|$ ===> increase in certainty
3) $[p_K(C) = r_K]$ less likely a priori ===> increase in certainty
4) $\forall c_j \in C$ $[p_K(c_j) = r_K]$ more likely ===> increase in certainty
5) $\forall C' \in C^S$, C' .NOT.= C $[p_K(C') \neq r_K]$
   $\cup C'$ less likely (union of all such sets)
   ===> increase in certainty
6)
7)
8) C more likely ===> increase in certainty

APPLICATION

Given that:

[classify physics research(countries with nuclear weapons) = true]
and [import quantities of uranium or plutonium
(countries with nuclear weapons) = true]

We can generate the inference rules:

[classify physics research(countries with nuclear weapons) = true]
[classify physics research(country X) = true]
===> country X countries with nuclear weapons, and
[import quantities of uranium or plutonium
(countries with nuclear weapons) = true]
[import quantities of uranium or plutonium(country X) = true]
===> country X countries with nuclear weapons

which we may then apply to:

[classify physics research(country X) = true], and
[import quantities of uranium or plutonium(country X) = true]

To infer that:

Country X is a country with nuclear weapons.

ANALOGY

If a property is in a given range for one case, and if a second matches the first on many properties, then infer that the property is in the same range for the second case.

EXAMPLE

This tank is like another tank the Russians made which is vulnerable to attack from above, so this one may be vulnerable as well.

CONDITIONS THAT INCREASE CERTAINTY

1) The more properties on which the two cases match.
2) The better the match on any property.
3) The more certain the property is in the given range for the one case.
4) The more likely the property is in the given range a priori.
5) The greater the semantic distance from the second case of the most similar case that has a contradictory property.

FORMULATION OF RULE

$t([p_K(c_1) = r_K])$ ::::>
$[p_K(c_1) = r_K] \wedge \forall p_\ell \in p \subseteq P \; ([p_\ell(c_1) = r_\ell] \wedge [p_\ell(c_2) = r_\ell])$
---> $[p_K(c_2) = r_K]$

Here, c_1 and c_2 represent particular cases, c_2 being a concept variable. The set, p is a subset of the properties of a concept for which the properties are known to be the same for particular cases, c_1 and c_2.

FORMULATION OF CERTAINTY

1) $\forall p_\ell \in p' \subseteq P \quad [p_\ell(c_1) = r_\ell] \wedge [p_\ell(c_2) = r_\ell]$,
$\forall p_\ell \in p'' \subseteq P \quad [p_\ell(c_1) = r_\ell] \wedge [p_\ell(c_2)$ .NOT.= $r_\ell]$
$|p'|$ greater wrt $|p''|$ ===> increase in certainty
2) $\forall p_\ell \in p \subseteq P \quad [p_\ell(c_1) = v_1] \wedge [p_\ell(c_2) = v_2]$, $v_1 \in r_\ell$, $v_2 \in r_\ell$,
v_1 more similar to v_2 ===> increase in certainty

3) $[p_K(c_1) = r_K]$ more certain ===> increase in certainty
4) $[p_K(c_2) = r_K]$ more likely a priori ===> increase
in certainty
5) $(\exists c_i\ \forall c_K,\ k \neq i,\ k \neq 2,\ ((\forall p_\ell \in p \subseteq P\ [p_\ell(c_i) = r_\ell],$
$\forall p_\ell \in \hat{p} \subseteq P\ [p_\ell(c_K) = r_\ell])$
$(|p| > |p|) \wedge$
$[p_K(c_2) = r_K] \wedge [p_K(c_i)$ .NOT.= $r_K]))$
$\wedge(\forall p_\ell \in p' \subseteq P\ [p_\ell(c_2) = r_\ell] \wedge [p_\ell(c_i) = r_\ell],$
$\forall p_\ell \in p'' \subseteq P\ [p_\ell(c_2) = r_\ell] \wedge [p_\ell(c_i)$ .NOT.= $r_\ell])$
$|p''|$ greater wrt $|p'|$ ===> increase in certainty

APPLICATION

Given that:

[Vulnerability(another tank) = from above]

We can generate the inference rule:

[Vulnerability(another tank) = from above]
$\forall p_\ell \in p \subseteq P$ $[p_\ell$(another tank) $= r_\ell] \wedge [p_\ell$(this tank) $= r_\ell]$
((that is, other properties of each tank have similar values))
---> [Vulnerability(this tank) = from above]

Which we may then apply to -another tank'-, a particular -another tank- with similar properties to -this tank-, to infer:

[Vulnerability(another tank') = from above]

GENERALIZATION (Degenerate proof by cases)

If a property is in a given range for some instance, then infer it is in that range for the set of which the instance is a member.

EXAMPLE

A MIG 1000 captured from the Russians cannot reach certain speeds at low altitude. Therefore, probably all the MIG 1000's cannot.

CONDITIONS THAT INCREASE CERTAINTY

1) The less variability in the set.
2) The more typical the instance is of the set.
3) The more certain the property is in the given range for the instance.
4) The greater the semantic distance from the set of the most similar case that has a contradictory property.
5) The more likely the property is in the given range a priori.

FORMULATION OF RULE

$t([p_K(c_i) = r_K]) \;::::>\; [p_K(c_i) = r_K] \;--->\; [p_K(C)\ r_K]$

C represents a concept variable, which is specialized to a particular case, c_i. The transformation produces an inference rule for property p allowing generalization from the particular case c_i to a more general concept, C.

FORMULATION OF CERTAINTY

1) $\forall p_\ell \in P \quad [p_\ell(C) = r_\ell]$
   $\Sigma |r_\ell|$ smaller wrt $|C|$ ===> increase in certainty
2)
3) $[p_K(c_i) = r_K]$ more certain ===> increase in certainty
4) $(\exists c_m \forall c_K,\ k \neq m,\ k \neq i,\ ((\forall p_\ell \in p \subseteq P \quad [p_\ell(c_m) = r_\ell],$
   $\forall p_\ell \in \hat{p} \subseteq P \quad [p_\ell(c_K) = r_\ell]) \wedge$
   $(|p| > |\hat{p}|) \wedge$
   $[p_K(c_i) = r_K] \wedge [p_\ell(c_m) \text{ .NOT.= } r_K]))$

$\wedge (\forall p_\ell \in P' \subseteq P_\ell\ [p_\ell(c_i) = r_\ell] \wedge [p_\ell(c_m) = r_\ell],$
$\forall p_\ell \in P'' \subseteq P\ \ [p_\ell(c_i) = r_\ell] \wedge [p_\ell(c_m)\ .NOT.= r_\ell])$
$|P''|$ greater wrt $|P'|$ ===> increase in certainty

5) $[p_K(C) = r_K]$ more likely a priori ===> increase in certainty

APPLICATION

Given that:

[capability(a MIG 1000 captured from the Russians) =
    cannot reach certain speeds at low altitude]

We can generate the inference rule:

[capability(a MIG 1000 captured from the Russians) =
    cannot reach certain speeds at low altitude]
  ---> [capability(a generalization of a MIG 1000 captured from
    the Russians) = cannot reach certain speeds at low altitude]

Which we can then apply to:

[a generalization of a MIG 1000 captured from the Russians =
    all MIG 1000's]

To infer:

[capability(all MIG 1000's) =
    cannot reach certain speeds at low altitude]

FUNCTIONAL DEDUCTION (or Calculation)

If a property (the dependent variable) depends on several factors (the independent variables), and if the values for the factors are in a given range for a particular case, then infer that for this case the property is in the range that follows from the values of the factors.

EXAMPLE

If the grain production of a country depends on the mean grain production for the country together with modifications for the distribution of rainfall and temperature during the growing season, then the production for a particular country can be estimated from these factors.

CONDITIONS THAT INCREASE CERTAINTY

1) The more the property depends on the factors for which the values are known.
2) The more certain that any of the factors are in the given range.
3) The more likely the value of the property is in the given range.

FORMULATION OF RULE

$$t([\ *\ ]\ [p_{i_1}(C) = r_{i_1}] \ldots [p_{i_n}(C) = r_{i_n}] \Longrightarrow [p_d(C) = r_d] \wedge [C = c_i])$$
$$:::\!> [p_{i_1}(c_i) = r_{i_1}] \ldots [p_{i_n}(c_i) = r_{i_n}] \longrightarrow [p_d(c_i) = r_d]$$

Here C represents a concept variable, which is specialized explicitly to a particular case by adding the conjunction $[C = c_i]$ to the antecedent formula. Note that functional deduction is the specialization rule implicitly used by functional abduction and functional analogy. Further, note that [*] implies that the functional dependency also includes independent variables whose values are not known.

FORMULATION OF CERTAINTY

1) $\forall p_{i_k} \in p' \subseteq p_i$ KNOWN(p_{i_k}, C),
$\forall p_{i_k} \in p'' \subseteq p_i$ .NOT.KNOWN(p_{i_k}, C)
((Note [*] = [$p''(C) = r''$]))

$|p'|$ greater wrt $|p''|$ ===> increase in certainty
2) $\forall i_k \in I' \subseteq I$
$[p_{i_k}(c_i) = r_{i_k}]$ more certain ===> increase in certainty
3) $[p_d(C) = r_d]$ more likely ===> increase in certainty

APPLICATION

Suppose for the above example, that a country's grain production depends on the mean grain production for the country, and the country's rainfall distribution and temperature during the growing season as per the following functional dependency, i.e. given this functional dependency:

```
Grain Production(country) =
    high [mean grain production(country) = high]
         [rainfall distribution(country) = moderate, high]
         [temperature(country) = moderate]  V
    moderate [mean grain production(country) = moderate]
         [rainfall distribution(country) = moderate]
         [temperature(country) = moderate]  V
    low [mean grain production(country) = moderate]
         [rainfall distribution(country) = moderate]
         [temperature(country) = low]
```

For a given country, let us say Botswana (that is [country = Botswana]), let us say that we know the following concept:

```
[mean grain production(Botswana) = moderate]
     [rainfall distribution(Botswana) = moderate]
```

Applying functional deduction to Grain production(country), based on the known variables, we get:

```
Grain production(Botswana) =
    high [mean grain production(Botswana) = high]
         [rainfall distribution(Botswana) = moderate, high]  V
    moderate [mean grain production(Botswana) = moderate]
         [rainfall distribution(Botswana) = moderate]  V
    low [mean grain production(Botswana) = moderate]
         [rainfall distribution(Botswana) = moderate]
```

After applying logical reduction methods to the above, the final result of functional deduction applied to Grain Production(Botswana), based on the known variables is:

```
high [mean grain production(Botswana) = high]
     [rainfall distribution(Botswana) = moderate, high]  V
moderate, low [mean grain production(Botswana) = moderate]
     [rainfall distribution(Botswana) = moderate]
```

Finally, we can infer:

[Grain production(Botswana) = moderate, low]

FUNCTIONAL ABDUCTION

If a property (the dependent variable) depends on several factor (the independent variables), and if the property is in a given rang for a particular case, then infer that any factor has a value in th range that leads to the given range on the dependent variable.

EXAMPLE

Whether Egypt attacks Israel depends on thinking that they wil not lose the Nile region, that Israel does not have atomic weapons and that Israel is not expecting an attack. Therefore when Egyp attacked Israel in 1967, they must have thought Israel did not hav atomic weapons.

CONDITIONS THAT INCREASE CERTAINTY

1) The fewer alternative functional dependencies that lead to the given range for the dependent variable.
2) The more dependent variables that are in the range that follows from the value of the factor(s).
3) The less likely that the property is in the given range a priori
4) The more likely that the property is in the given range for the values of the factors.
5) The less likely are alternative functions that lead to the given range for the dependent variable.
6) The less correlated are the dependent variables that are in the range that follows from the value of the factor(s).
7) The more correlated are the alternative functions that lead to the given range for the dependent variable.
8) The more likely are the particular values of the factor(s).

FORMULATION OF RULE

$$t([\ *\ ]\ [p_{ik}(C) = r_{ik}] ===> [p_d(C) = r_d])$$
$$::::> ([p_d(C) = r_d] ---> [p_{ik}(C) = r_{ik}])$$

where C represents a concept variable, which may be specified eithe in the antecedent formula or the inferred formula by the addition o the conjuction, [C = c_i], where c_i is an element is set c, th possible values of C. Note that functional deduction is required a a specialization operation to determine a value c_i for C, in orde to apply functional abduction to a particular case.

FORMULATION OF CERTAINTY
----------- -- ---------

1) $\forall p'_i \in P_i$, $[p'_{i_1}(C) = r'_{i_1}] \ldots [p'_{i_n}(C) = r'_{i_n}] \Longrightarrow$
   $[p_d(C) = r_d]$, $r'_{i_\ell} \in R'_i$
   $|P_i|$ smaller ===> increase in certainty
2) $\forall c_j \in C' \subseteq C$ $[p_{i_k}(c_j) = r_{i_k}]$
   $|C'|$ greater ===> increase in certainty
3) $[p_d(C) = r_d]$ less likely a priori ===> increase in certainty
4) $[p_{i_1}(c_i) = v_{i_1}] \ldots [p_{i_n}(c_i) = v_{i_n}] \Longrightarrow [p_d(c_i) = r_d]$
   more likely ===> increase in certainty
5) $\forall p'_i \in P_i$, $p'_i \neq p_i$
   $[p'_{i_1}(C) = r'_{i_1}] \ldots [p'_{i_n}(C) = r'_{i_n}] \Longrightarrow [p_d(C) = r_d]$
   less likely ===> increase in certainty
6)
7)
8) $\forall p_{i_\ell} \in P_i \subset P$ $[p_{i_\ell}(c_i) = v_{i_\ell}]$
   v_{i_ℓ} more likely ===> increase in certainty

APPLICATION

Given that:

([think will lose Nile region(Egypt) = false]
   [think Israel expecting attack(Egypt) = false]
   [think Israel has atomic weapons(Egypt) = false]
   ===> [attacks Israel(Egypt) = true])

We can generate the inference rule:

[attacks Israel(Egypt) = true]
   ---> [think Israel has atomic weapons(Egypt) = false

which we may then apply to:

[Egypt = Egypt in 1967]

Using functional deduction to obtain the particular inference rule:

[attacks Israel(Egypt in 1967) = true]
   ---> [think Israel has atomic weapons(Egypt in 1967) = false]

To infer:

[think Israel has atomic weapons(Egypt in 1967) = false]

Note that by applying functional abduction, we can also derive rul
which allow us to infer:

[think will lose Nile region(Egypt in 1967) = false] and
[think Israel expecting attack(Egypt in 1967) = false]

FUNCTIONAL ANALOGY

If a property (the dependent variable) depends on a number factors (the independent variables), and if one case matches anothe case on these factors, and if the value of the property for one cas is in a given range, then infer that the value of the property fo the other case is in the given range.

EXAMPLE

When Commander X was making little headway along a particula front, and the dry weather season had only two months to run, h launched an attack by sea at a position that was unprotected. Sinc he is in the same position now, we should expect another surpris attack from him.

CONDITIONS THAT INCREASE CERTAINTY

1) The more factors on which the two cases match.
2) The greater the dependency on any factors on which the two cases match.
3) The better the match on any factor.
4) The more certain the property is in the given range for the one case.
5) The more likely the value of the property is in the given range a priori.
6) The more certain the factors have the given values in the two cases.
7) The greater the dependency on those factors that match best.

FORMULATION OF RULE

$$(\bar{C} +]\ [p_{i_1}(C) = r_{i_1}] \ldots [p_{i_n}(C) = r_{i_n}] \Longrightarrow [p_d(C) = r_d])$$
$$:::: > [p_{i_1}(C_1) = r_{i_n}]\ [p_{i_1}(C_2) = r_{i_1}] \ldots [p_{i_n}(C_1) = r_{i_n}]$$
$$[p_{i_n}(C_2) = r_{i_n}]\ [p_d(C_1) = r_d] \longrightarrow [p_d(C_2) = r_d]$$

Here C_1, and C_2 represent particular cases with C being a concept variable which may be specified in the generated formula by the conjunction $[C_2 = c_2]$, where c_2 is an element of set c, the possible values of C_2. Note that functional deduction is require as a specialization operation to determine the value c_2 for C_2. Note also that the form of the functional dependency is not know for the general case, as it was for functional deduction and functional abduction, but just that such a dependency exists. Knowing the form for a particular case, however, permits the

inference of the form for a similar case by functional analo
Here [*] represents properties whose values are either not kn
for C_1, not known for C_2, or not in the same range for C_1 and C_2.

FORMULATION OF CERTAINTY

1) $\forall p_{i\ell} \in p'_i \subset p_i$ $[p_{i\ell}(c_1) = r_{i\ell}] \wedge [p_{i\ell}(c_2) = r_{i\ell}]$,
$\forall p_{i\ell} \in p''_i \subset p_i$ $[p_{i\ell}(c_1) = r_{i\ell}] \wedge [p_{i\ell}(c_2)$ .NOT.= $r_{i\ell}]$
$|p'|$ greater wrt $|p''|$ ===> increase in certainty
2) $\forall p_{i\ell} \in p'_i \subset p_i$ $[p_{i\ell}(c_1) = r_{i\ell}] \wedge [p_{i\ell}(c_2) = r_{i\ell}] \wedge$
$([p\ (C) = r\ ]$ ===> $[p\ (C) = r\ ]$ more likely$)$
===> increase in certainty
3) $\forall p_{i\ell} \in p_i$ $[p_{i\ell}(c_1) = v_1] \wedge [p_{i\ell}(c_2) = v_2]$ $v_1, v_2 \in r_{i\ell}$
v_1 more similar to v_2 ===> increase in certainty
4) $[p_d(c_1) = r_d]$ more certain ===> increase in certainty
5) $[p_d(c_2) = r_d]$ more likely a priori ===> increase in certainty
6) $\forall p_{i\ell} \in p_i$ $[p_{i\ell}(c_1) = v_1]$ more certain, $[p_{i\ell}(c_2) = v_{i\ell}]$
more certain, $v_1, v_2 \in r_{i\ell}$ ===> increase in certainty
7) $\forall p_{i\ell} \in p_i$ $[p_{i\ell}(c_1) = v_1] \wedge [p_{i\ell}(c_2) = v_2] \wedge v_1$ very similar
to v_2, $v_1, v_2 \in r_{i\ell}$
$[p_{i\ell}(C) = r_{i\ell}]$ ===> $[p_d(C) = r_d]$ more likely
===> increase in certainty

APPLICATION

Given that:

[making little headway along front(commander X once before) = true]
[dry weather season left(commander X once before) = 2 months]
---> [launch attack at(Commander X once before)
= unprotected position]

We can generate the inference rule:

[making little headway along front(commander X once before) = true]
[making little headway along front(commander X another time)
= true]
[dry weather season left(commander X once before) = 2 months]
[dry weather season left(commander X another time) = 2 months]
[launch attack at(commander x once before)
= unprotected position]
---> [launch attack at(commander X another time)
= unprotected position]

which we may then apply to:

```
[commander X another time = commander X this time]
```

Using function deduction to obtain the particular inference rule:

```
[making little headway along front(commander X once before) = true]
   [making little headway along front(commander X this time)
      = true]
   [dry weather season left(commander X once before) = 2 months]
   [dry weather season left(commander X this time) = 2 months]
   [launch attack at(commander X once before)
      = unprotected position]
   ---> [launch attack at(commander X this time)
      = unprotected position]
```

which we apply to:

```
[making little headway along front(commander X this time) = true]
   [dry weather season left(commander X this time) = 2 months]
```

To infer:

```
[launch attack at(commander X this time) = unprotected position]
```

INFERENCES ON META-KNOWLEDGE

LACK-OF-KNOWLEDGE-INFERENCE (Meta Modus Tollens)

If a person would know about a property for a given case if it wer in a given range and if the person does not know about the property then infer that the property is NOT in the given range for tha case.

EXAMPLE

If a particular road has been bombed, this information would have been reported by out observers in the area. It has not been, so therefore the road is probably in good shape.

CONDITIONS THAT INCREASE CERTAINTY

1) The more important the particular case.
2) The less likely the property is in the given range.
3) The more information stored about the given case.
4) The more similar properties stored about the given case.
5) The more important is the property to other similar cases.
6) The more important the given property.
7) The more information stored about the given property.
8) The more similar cases stored with the given property.
9) The more important the case to other similar properties.

FORMULATION OF RULE

(([p_K(C) = r_K] ===> KNOWN(p_K, C)) ::::> .NOT.KNOWN(p_K, C)
    ---> [p_K(C) .NOT.= r_K]

Here C represents a concept variable, which may be specified by the addition of the conjunction, [C = c_i], where c is an element of set c, the possible values of C. Note that functional deduction must first be performed on the antecedent formula with [C = c], in order to specialize C. KNOWN(p_K, C) is a function returning true if property p_K is know for case C, and false otherwise, .NOT.KNOWN(p_K, C) is the complementary function.

FORMULATION OF CERTAINTY

1) c_i more important in C^S ===> increase in certainty
2) $[p_K(C) = r_K]$ less likely ===> increase in certainty
3) $\forall p_\ell \in p \subseteq P$ KNOWN(p_ℓ, c_i)
   $|p|$ greater ===> increase in certainty
4) $\forall p_\ell \in p \subseteq P$ KNOWN(p_ℓ, c_i)
   p_ℓ similar to p_K ===> increase in certainty
5) $\forall c_j \in c \subseteq C$ ($\forall p_\ell \in p' \subseteq P$ KNOWN(p_ℓ, c_i) $\wedge$
   $[p_\ell(c_j) = r_\ell] \wedge [p_\ell(c_j) = r_\ell]$,
   $\forall p_\ell \in p'' \subseteq P$ KNOWN(p_ℓ, c_i) $\wedge$
   $[p_\ell(c_j) = r_\ell] \wedge [p_\ell(c_j)$ .NOT.= $r_\ell]$
   $|p'|$ greater wrt $|p''|$)
   p_K more important to c_j ===> increase in certainty
6) p_K more important to c_i ===> increase in certainty
7) $\forall c_j \in c' \subseteq C^S$ KNOWN(p_K, c_j)
   $\forall c_j \in c'' \subseteq C^S$ .NOT.KNOWN(p_K, c_j)
   $|c'|$ greater wrt $|c''|$ ===> increase in certainty
8) $\forall c_j \in c' \subseteq C$ KNOWN(p_K, c_j)
   ($\forall p_\ell \in p' \subseteq P$ $[p_\ell(c_i) = r_\ell] \wedge [p_\ell(c_j) = r_\ell]$,
   $\forall p_\ell \in p'' \subseteq P$ $[p_\ell(c_i) = r_\ell] \wedge [p_\ell(c_j)$ .NOT= $r_\ell]$)
   $|p'|$ large wrt $|p''|$,
   $\forall c_j \in c'' \subseteq C$ .NOT.KNOWN(p_K, c_j)
   ($\forall p_\ell \in p' \subseteq P$ $[p_\ell(c_i) = r_\ell] \wedge [p_\ell(c_j) = r_\ell]$,
   $\forall p_\ell \in p'' \subseteq P$ $[p_\ell(c_i) = r_\ell] \wedge [p_\ell(c_j)$ .NOT= $r_\ell]$)
   $|p'|$ large wrt $|p''|$,
   $|c'|$ greater wrt $|c''|$ ===> increase in certainty
9) c_i more important to properties similar to p_K
   ===> increase in certainty

APPLICATION

Given that:

[Condition(road) = bombed] ===> KNOWN(condition, road)

We can generate the inference rule:

.NOT.KNOWN(condition, road) ---> [condition(road) = .NOT.bombed]

Which we may then apply to:

[road = particular road]

Using functional deduction to obtain the particular inference rule:

```
.NOT.KNOWN(condition, particular road)
   ---> [condition(particular road) = .NOT.bombed]
```

To infer:

```
[condition(particular road) = .NOT.bombed]
```

REFERENCES

1. Polya, G., "Patterns of Plausible Inference", Princeton University Press, Princeton, New Jersey, 1968.

2. Collins, A., "Fragments of a Theory of Human Plausible Reasoning", Proceedings of TINLAP 2, July 1978.

3. Collins, A., A listing of 13 Plausible Inference Rules, unpublished.