Mental MetaLogic: A New Paradigm in Psychology of Reasoning

Yingrui Yang & Selmer Bringsjord  
yangyri@rpi.edu • selmer@rpi.edu  
Dept. of Philosophy, Psychology & Cognitive Science (S.B. & Y.Y.)  
The Minds & Machines Laboratory (S.B. & Y.Y.)  
Department of Computer Science (S.B.)  
Rensselaer Polytechnic Institute  
Troy NY USA

Abstract

We introduce a new theory of reasoning, Mental MetaLogic (MML), which unifies the mental logic and mental models theories currently deadlocked in opposition to each other. MML, as we explain, is based on the “psychologizing” of all of symbolic logic — which includes a syntactic component (proof theory) with which mental logic is associated, a semantic component (model theory and its relatives) with which mental models is associated, and a “meta” component (metatheory or metalogic), hitherto largely ignored in the psychology of reasoning. We briefly summarize two samples of MML-based empirical/experimental work.

1 Introduction

Few would dispute the claim that logical reasoning is a fundamental human ability. Smith can’t find his car keys, but knows both that he last had them minutes ago in the kitchen, and that he has only been in one other room since then. After scouring every millimeter of the other room, he reenters the kitchen and, with confidence, pokes around and finds them. His confidence is the result of deductive reasoning, that much we know. But what is deductive reasoning? In the last three decades, the psychology of reasoning has produced countless papers in psychological journals (such as Psychological Review, Science, and Behavioral & Brain Sciences) aimed at providing an answer to this question, and aimed thereby at providing a scientific explanation of what Smith has just done. Alas, the answer and explanation have never arrived. The reason is that if one had to capture the psychology of reasoning in a word or two, the first choice would doubtless be “The Controversy.”

The Controversy is well-known to many of our readers. In brief, it arises from two competing theories for modeling and explaining ordinary human reasoning (especially deductive reasoning). On the one hand, mental logic theory claims that logically untrained humans reason using domain-independent inference rules or schemas (e.g., that from $p \lor q$ and $\neg p$ one can infer to $q$; proponents of mental logic would say that something like this particular rule is operative in the case of our hypothetical Smith). On the other hand, mental model theory holds that ordinary humans reason using not abstract syntactic rules, but semantic possibilities they in some sense imagine. Scientists on both sides are incontestably seminal and prolific, and the content offered by each is powerful. But the two sides stand vehemently apart, and as a result the science of reasoning is, in a real sense, fundamentally paralyzed. MML will end this paralysis.

MML is a meta-theory because it is inspired by modern symbolic logic, which is primarily a meta-inquiry, as it is devoted to the mathematical study of mathematical and logical systems.1 And although MML is a psychological theory, to present it we need a quick review of the three main components of symbolic logic.

In broad strokes, modern symbolic logic has three main components: one is syntactic, one is semantic, and one is meta-theoretical. The syntactic component includes specification of the alphabet of a given logical system, the grammar for building well-formed formulas (wffs) from this alphabet, and a proof theory that precisely describes how and when one formula can be proved from a set of formulas. The semantic component includes a precise account of the conditions under which a formula in a given system is true or false. The meta-theoretical component includes theorems, conjectures, and hypotheses concerning the syntactic component, the semantic component, and connections between these two components. (E.g., meta-theory would include proofs of such things as that all arguments whose conclusions are a (semantic) consequence of their premises can be proved in step-by-step, syntactic fashion.)

---

1Tellingly, mathematical logic, a part of symbolic logic, is sometimes called ‘meta-mathematics.’
2 The Components & ML, MM, MML

Now, using the review provided in the previous section, we characterize mental logic, mental models, and Mental MetaLogic. This characterization is summed up in Figure 1.

The representational system of current mental logic theory can be viewed as (at least to some degree) a psychological selection from the syntactic components of systems studied in modern symbolic logic. For example, in mental logic *modus ponens* is often selected as a schema while *modus tollens* is not; yet both are valid inferences in most standard logical systems. Another example can be found right at the heart of Lance Rips’ system of mental logic, PSYCOP, set out in (Rips 1994), for this system includes conditional proof (p. 116), but not the rule which sanctions passing from the denial of a conditional to the truth of this conditional’s antecedent (pp. 125–126).

In parallel to the relationship between mental logic and the syntactic component of symbolic logic, mental model theory is related to the semantic side of symbolic logic. For example, consider the truth table for conditionals in the propositional calculus, shown in Figure 2. The representational system of the mental model theory consists of a number of sets of initial mental models selected from the semantic possibilities. Because each proposition can have two truth-values, true or false, the truth table as a truth function for → has four semantic (truth) possibilities. When $p$ is true and $q$ is false, the conditional is false. By the truth principle of mental model theory, reasoners represent only what is true but not what is false, due to limited working memory. So this possibility (row 3) is excluded from the set of possibilities. Proponents of MM also predict that ordinary reasoners will not reason from a false antecedent, thus the two semantic possibilities where $p$ is false (rows 4 and 5) are not represented explicitly, but rather are represented implicitly by making a mental footnote as “…” Thus, the only explicit model left is the case in which both $p$ and $q$ are true (row 2). Hence, by using the presence/absence of a mental token to substitute the truth value accordingly, the initial model set for the form of conditionals consists of one explicit and one implicit mental model, which can be represented as $p \quad \ldots \quad q$

As we have seen earlier, the meta-theoretical component of modern symbolic logic covers formal properties (e.g., soundness and completeness) that bridge the syntactic and semantic components of logical systems. In selecting from either but not both of the syntactic or semantic components of logical systems, psychologists have, in a sense, broken the bridges built by logicians and mathematicians. Mental logic theory and mental model theory became incompatible from the perspective of symbolic logic: the former is explicitly and uncompromisingly syntactic, while the latter is explicitly and uncompromisingly semantic. Mental MetaLogic aims to bridge between these two theories.

3 Glimpse of MML in Action

The following two samples indicate how it’s possible for a psychological metatheory to resolve two major recent challenging issues in the domain of quantified predicate reasoning.

**Sample 1.** People can solve many problems, but of course they don’t find all these problems
equally difficult. In other words, we can distinguish between two levels of cognitive routines in reasoning: Reasoners can solve certain kinds of inference problems errorlessly, and they are able to perceive systematically the difference between these problems in terms of their difficulty. Mental logic theory proposes a set of inference schemas, and predicts

P1 Reasoners can apply them in reasoning errorlessly, and
P2 Each schema has its particular difficulty.

For example, consider a pair of verbal instantiations of two schemas:

(1) All the children got some red beads. Therefore, all the girls got red beads.

(2) All the beads are either red or blue. The red beads are square. The blue beads are round. Therefore, all the beads are either round or square.

People can apply both schemas errorlessly, but we need to make sure that (2) is harder than (1). Now consider a third problem (3):

All the children found some red beads.
The red beads were either round or square.
The round beads were plastic.
The square beads were wooden.

Did all the girls find either plastic or wooden beads?

This problem is soluble by using the appropriate variations of the schemas in (1) and (2). Yang, Braine & O’Brien (1998) conducted a large-scale project to examine the mental predicate logic devised by Braine (1998). About 130 inference problems like (3) were constructed and all the problems were soluble by using 10 direct reasoning schemas like (1) and (2). The participants (N=180, all experiments were individually administered) were instructed to solve each problem first and then to rate its difficulty on a 7-point scale. As predicted, the overall error rate (of more than 13,000 responses) was lower than 3%. Thus, the introspective difficulty ratings could be used to generate the difficulty-weight for each schema. A parametrical model was used to generate schema-weights by using a least-square method. The theory predicts the mean difficulty of an inference problem by using the sum of difficulty-weights of those schemas included in the proposed solution to solve that problem. As to results, the cross-validation tests (i.e., using the set of weights generated from one data-set to predict the mean difficulty ratings of any other data-sets; total of 9 samples were used) show that the correlation was reliably as high as 0.93, accounting for more than 80% of the variance. This result provides direct empirical evidence supporting mental logic theory, and challenges the mental model theory because (i) current mental model theory cannot represent some of the schemas (e.g., a schema involving a quantified conjunction), and (ii) the weights were generated from empirical data. The question is, how could mental model theory account for the schema weights?

Mental MetaLogic responds this challenge in two steps. The first step is to modify the mental model representation in order to represent all the schemas. For example, the form $\forall x (\text{IF } P x \text{ THEN } (F x \text{ OR } G x))$ is represented by the model $[p]^f_g$

where $[ ]$ indicates the universal quantifier, and together the superscript $f$ and the subscript $g$ indicates $f \lor g$. This introduces a logical structure into a single model, which is different from the original version of mental model representation. For the second step, as an example, consider the quantified modus ponens:

All the As are Bs. All the Bs are Cs.

All the As are Cs.

Its model representation is as follows $[a] b a[b] c$

Here the left-box indicates the model for the first premise, and the next box indicates the result of integrating the information from the second premise with the previous model. Now, one can easily count that the number of the mental tokens used in this model representation is 5. Yang (n.d.) found that the number of the mental tokens included in a consistently modified mental model representation of this type is a very accurate predictor for the schema weights. That is, for each schema examined, its difficulty weight equals the number of the mental tokens included in its minimum mental model representation (divided by 10). The correlation between the two sets of weights (predicted by the mental model method and generated by the mental logical parametrical method) is almost perfect ($r = .99$).

Sample 2. People can make many correct inferences if they use the schematic representation proposed by mental logic theory; there is simply no questioning this. People can also get many inferences right if they use mental model
representations; however, mental model theory also accommodates illusory inferences, which seem very compelling but are fallacious. Consider the following examples:

Only one of the following statements is true:
– Some of the plastic beads are not red, or
– None of the plastic beads are red.

Problem 1 Is it possible that none of the red beads are plastic?

Problem 2 Is it possible that at least some of the red beads are plastic?

As we said earlier, Problem 1 is predicted by mental model theory as an illusory problem (i.e., the correct answer is “No” but most people would say “Yes”), and Problem 2 is predicted to be a control problem (i.e., most people can get the correct answer, “Yes”). Johnson-Laird and his colleagues have published, at last count, a dozen of papers on this topic, including the recent Science paper (P. N. Johnson-Laird, Girotto & Legrenzi 2000). These results have caused something arguably approaching a crisis for mental logic theory, because current mental logic theory cannot account for both the controls and the illusions (Yang & Johnson-Laird 2000). But here is how MML works in a series of steps toward resolving the crisis. First, MML analyzes the difficulties facing mental logic theory: (1) Johnson-Laird’s illusory phenomenon requires an extension of mental logic theory to deal with possibility, which the current theory does not handle. (2) The phenomenon requires mental logic theory to modify the Principle of Truth via the mental model claim that people often represent only what is true but not what is false. This is difficult because the principle concerns the truth-values (true or false), but any current schema is used only “at its face value.” (3) The phenomenon requires allowing some invalid schemas. It is conceptually difficult to swallow this from the classical metatheoretical point of view, because in classical logic and mathematics everything follows from a contradiction.

The second step involves a revision of the current version of the theory. For difficulty (3), we say invalid schemas should be allowed, because as psychologists spanning The Controversy, we have after all agreed that, in (untrained) human reasoning, nothing follows from contradiction. For (2), we revise the original definition of the inference schema due to Braine (1998) by adding one clause as italicized just below.

Notice that Braine formulated inference schemas with certain semantic elements. For example, because his predicate logic is domain-specific, each quantified statement in a schema is with a given domain, which belongs to formal semantics in classical logic. Then we can define the Understanding Conditional as:

\[ p \Rightarrow x \] is read as \( p \) understandingly implies \( x \), meaning \( x \) is a semantic consequence of \( p \).

Accordingly, \( p \Rightarrow x \) should be read as If-Then when \( x \) is given. For the third step, we can now define two illusory schemas as below. First we have

\[ (p \rightarrow x) \text{ori} (q \rightarrow x); p \text{ore} q; \therefore x \]

where “ori” denotes inclusive disjunction and “ore” exclusive disjunction. The variable \( x \) may include negation or some modal operator, or both. For instance, here is a specific version of the schema accounting for illusory inferences of impossibility:

\[ (p \Rightarrow \neg \diamond s) \text{ori} (q \Rightarrow \neg \diamond s); p \text{ore} q; \therefore \neg \diamond s \]

Handled in this way, illusory schemas accommodate the Principle of Truth, and it is general enough for the proof-theoretic approach to account for the published illusory inferences caused mainly by exclusive disjunction, and thus it should also predict the same type of illusory inferences with compound statements. Some other illusory schemas would of course need to be developed to account for other possible type of illusory inferences (and corresponding work is required for the full adaptation of mental model theory as well).

4 Acknowledgments

Our MML research is currently supported by the National Science Foundation and the Educational Testing Service. We’re grateful for this indispensable support. We’re also greatly indebted to Phil Johnson-Laird for countless objections, insights, and suggestions.

References


Yang, Y. (n.d.), How mental model theory can predict quantified schema-weights.
