ELSEVIER

Cognitive Science 25 (2001) 839–901

COGNITIVE SCIENCE

http://www.elsevier.com/locate/cogsci

Model theory of deduction: a unified computational approach

Bruno G. Bara[a], Monica Bucciarelli[a,*], Vincenzo Lombardo[a,b]

[a]Centro di Scienza Cognitiva, Università di Torino, via Po 14, 10123 Torino, Italy
[b]Dipartimento di Informatica, Università di Torino, 10123 Torino, Italy

Abstract

One of the most debated questions in psychology and cognitive science is the nature and the functioning of the mental processes involved in deductive reasoning. However, all existing theories refer to a specific deductive domain, like syllogistic, propositional or relational reasoning.

Our goal is to unify the main types of deductive reasoning into a single set of basic procedures. In particular, we bring together the microtheories developed from a mental models perspective in a single theory, for which we provide a formal foundation. We validate the theory through a computational model (UNICORE) which allows fine-grained predictions of subjects' performance in different reasoning domains.

The performance of the model is tested against the performance of experimental subjects—as reported in the relevant literature—in the three areas of syllogistic, relational and propositional reasoning. The computational model proves to be a satisfactory artificial subject, reproducing both correct and erroneous performance of the human subjects. Moreover, we introduce a developmental trend in the program, in order to simulate the performance of subjects of different ages, ranging from children (3–6) to adolescents (8–12) to adults (>21). The simulation model performs similarly to the subjects of different ages.

Our conclusion is that the validity of the mental model approach is confirmed for the deductive reasoning domain, and that it is possible to devise a unique mechanism able to deal with the specific subareas. The proposed computational model (UNICORE) represents such a unifying structure.

Keywords: Mental models; Deduction; Computational model; Development

* Corresponding author. Tel.: +011-6703038; fax: +011-8159039.

E-mail addresses: monica@psych.unito.it (M. Bucciarelli), bara@psych.unito.it (B.G. Bara), vincenzo@di.unito.it (V. Lombardo).

0364-0213/01/$ – see front matter
PII: S0364-0213(01)00056-8

1. Introduction

Deduction is a systematic process whose goal is to draw a valid consequence from a series of premises. It requires one to consider the premises as true and to infer what conclusion, if any, follows. By definition, a *valid* deduction yields a conclusion that must be true given that the premises are true.

But, human deductions are not always valid. Indeed, the inferential process can be affected by several semantic factors like, for instance, the problem's content or the reasoner's beliefs. For this reason, logic is not an account of human deductive reasoning (see, e.g., Harman, 1986). Further, logic is a wrong normative theory because it permits inferences that naive individuals are not liable to draw (Devlin, 1997). A psychological theory of deduction has to account for plausible cognitive procedures, and to predict both the valid and the invalid deductions that humans draw.

In this paper we deal with three types of deduction: syllogistic, relational and propositional reasoning. They involve, respectively, reasoning from quantified assertions (e.g., 'All of the artists are beekeepers; All of the beekeepers are chemists'), premises containing relations (e.g., 'The bicycle is bigger than the doll; The doll is bigger than the pen'), and propositions involving connectives (e.g., 'If Ann is at the party, then Ben is there; Ann is at the party'). Mental Model Theory (MMT: Johnson-Laird, 1983; Johnson-Laird & Byrne, 1991) can explain human performance in all these three reasoning domains. As a matter of fact, the theory extends further over different areas, and most of them are currently covered by specific computational models independently developed.

Although earlier work (reviewed later) suggests that MMT can explain human performance in each reasoning domain separately, we show that one unified computational model can be at work in all the three domains. Such a model enhances the coherence of the theory, which is intended to account for any sort of deduction. Moreover, in formulating such a unified model, we do not sacrifice the explanatory adequacy for any of the individual reasoning domains; our unified model is as good as other MMT models developed specifically for each single area. The existence of such a computational model and its experimental performance demonstrate that a unified deductive mechanism is a viable hypothesis. There is no need to postulate specific mental processes associated with different deductive domains.

Our model assumes – in accordance with MMT – that any kind of reasoning consists of five main processes: construction of mental models of the premises, integration of mental models, formulation of conclusions consistent with the integrated models, falsification of conclusions, and production of (linguistic or motor) responses. To accomplish this, we assume that there exist a set of procedures which is common to any kind of deduction, and which is part of the competence of the human system. The implementation of such procedures on a computer requires the definition of an ontology of mental model objects, together with the definition of the basic abilities underlying deductive reasoning. Our ontology provides the first formal foundation for MMT. Hitherto it has never been made explicit within the mental model framework.

The unified theory we propose gains strength by virtue of wide scope and explanative parsimony along three dimensions: it applies across reasoning domains, it accounts for both competence and performance, it has a relevance from a developmental perspective. The cross-domain validity of the model relies on the fact that predictions of subjects' responses are grounded on a single basic mechanism, whose functioning can be affected by cognitive constraints.

Second, the model reflects the competence/performance dichotomy: it predicts valid conclusions when it accesses and correctly executes the procedures (competence level); it predicts specific erroneous responses when the constraints on the cognitive resources affect the execution of the procedures (performance level).

Finally, the model is relevant from a developmental perspective, because it allows us to predict different responses by subjects belonging to different age groups. Indeed, we follow the methodology of developmental cognitive science (Bara, 1995; Karmiloff-Smith, 1992). This means studying a mental process not only in its mature state, an approach which exclusively takes into account the final stage, but rather concentrating on how a given function develops from the infant to the child, through adolescence, up to adulthood, and finally decays in old age. Therefore, we do not focus on real-time processing within steady-state systems, but on changes in the ability to reason.

The unified model is validated by a comparison of its predictions with the experimental data in the literature, and by a new experiment we carried out for propositional reasoning. The results show that our computational model reproduces both the competence and the performance level of subjects belonging to different age groups. The existence of the software program demonstrates the sufficiency of the unified hypothesis of deductive reasoning.

The paper is in 9 sections: after a survey of the main theories of reasoning and their limits (2), we present the ontology of mental models (3) and a computational theory of reasoning (4), with its predictions (5). In section 6 we analyze syllogisms, then relational (7) and propositional (8) reasoning. A general discussion and the conclusions are in section 9.

2. A review of the current theories of reasoning

In reviewing the huge psychological literature about deductive reasoning, it is indispensable to fix some criteria to evaluate the different contributions. We have selected three dimensions, according to which we examine the literature:[1]

1. Can the theory scale up to explain multiple domains of reasoning, or is it restricted to a single domain? *Ceteris paribus,* a global and parsimonious theory is preferred over local microtheories. In principle, this criterion could be extended beyond deduction; namely, does the theory elucidate other sorts of reasoning, for example, reasoning inductively, and even other sorts of cognition, for example, the comprehension of discourse?
2. Does the theory account for the differences in competence and performance? A theory should not only predict the correct responses, but also the incorrect ones. Systematic errors have to be both predicted and explained, by means of the basic tenets of the theory.
3. Can the theory explain a developmental trend? It may be awkward to compare two steady-state theories, but if one of the two is able to predict the increase of performance from children to adults, then it is definitely better grounded.

We sketch now the theories developed as accounts of syllogistic, relational and propositional reasoning. These are the deductive domains analyzed the most in the psychological literature. Finally, we introduce the relevant attempts at unified theories for deductive reasoning.

Historically, psychology researchers have mainly focused on *syllogistic inference*, which, from Aristotle until Frege, has been the main topic of logic.[2] Psychologists have never proposed a complete theory of syllogistic reasoning based on formal rules of inference, because differences in difficulty among syllogisms are not explained by the length of their formal derivations. For a notable exception see Rips (1994) below. Some theories have been based on Euler circles or equivalent strings of symbols (e.g., Guyote & Sternberg, 1981), but the number of Euler representations required by syllogisms does not correlate with their difficulty. Indeed, an easy syllogism solved by 100% of adult subjects (Some of the A are B, All of the B are C; Therefore, Some of the A are C) would require 16 possible combinations of the representations of its premises, whereas subjects are able to consider at most four of the possible combinations (see Johnson-Laird & Bara, 1984). Recently, a new method based on Euler circles, but avoiding combinatorial explosion, has been proposed by Stenning and Oberlander (1995), who augment the graphical representation with a propositional notation. This increases the expressiveness of Euler circles, which become as powerful as MMT. Nevertheless, none of these theories comply with two of the above mentioned criteria: they cannot be extended to account for inferences other than syllogistic, and none of them has the potential to explain developmental trends.

Propositional reasoning has only recently become of interest to psychologists.[3] Many theorists have proposed models based on formal rules of inference that are domain independent. For instance, Piaget and his colleagues (Inhelder & Piaget, 1958; Piaget, 1953) claim that people can deal with propositional deductions because they possess a mental logic; propositional logic, they argue, would develop by the early teens. Other theorists have proposed natural deduction systems, where rules are claimed to have more psychological plausibility than standard logic (Braine, 1978; Braine, 1990; Braine, 1998; Braine & Rumain, 1983; Rips, 1990). Cheng, Holyoak and colleagues postulate, for the case of reasoning with the conditional connective, the existence of abstract knowledge structures such as *causation*, *obligation* and *permission* (Cheng & Holyoak, 1985; Cheng, Holyoak, Nisbett & Oliver, 1986). As far as conditional reasoning is concerned, it has also been claimed that there exist cognitive procedures specialized for reasoning about social situations and specific types of adaptive problems (Cosmides & Tooby, 1994; Gigerenzer & Hug, 1992; Platt & Griggs, 1993). In a radically alternative view, Pollard (1981) and Griggs and Cox (1982) propose that the specific experiences encoded in the memory of the system are the only determinants of such reasoning processes. All these theories fail to comply with our criteria. First, they are limited to propositional reasoning; in fact, some of them only account for the ability to reason with a single connective. Second, the theories fail to supply either a comprehensive account of correct performance or errors in reasoning. On the one hand, pragmatic schemas and memories for specific events do not explain how people can draw - in principle - correct conclusions independently of the content of the reasoning problem. On the other hand, mental logic theories have difficulty in explaining errors (Johnson-Laird, 1997). As a consequence, they are disarmed toward explaining developmental trends in reasoning.

In the case of *relational reasoning*, interest has focused mainly on three-term series problems[4] and two-dimensional spatial deductions.[5] For three-term series problems, meaning postulates (Bar-Hillel, 1967) or linguistic representations (Clark, 1969) have been proposed. It is though impossible to choose between the two approaches since they do not make different predictions about problem difficulty. Several proposals have been advanced in order to explain what kind of mental representations allow people to draw spatial inferences. The dispute is mainly between mental representations in a propositional form (e.g., Pylyshyn, 1973), or analogical representations (e.g., Byrne & Johnson-Laird, 1989). Theories of propositional representations claim that there exist formal rules of inference for two-dimensional spatial deductions (see, for instance, the set proposed by Hagert, 1984), but derivations using such rules again do not account for problem difficulty. Experimental data are consistent with the claims of the analogical theories that people reason by imaging the state of affairs described in the premises (Byrne & Johnson-Laird, 1989). All such theories, except the analogical one, are limited to relational reasoning. Further, they have been proved to account neither for differences between competence and performance, nor for differences in performance across age groups.

From a different perspective, Evans (1989) and Evans and Over (1996) have proposed to analyze actual reasoning performance rather than the competence mechanisms for reasoning. They propose a dual process theory of thinking in which tacit and parallel processes of thought - such as biases - combine with explicit and sequential processes in determining our course of actions. Biases are systematic tendencies either to account for factors irrelevant to the task at hand, or to ignore relevant factors. They are claimed to affect the reasoning process across deductive tasks, but they are not sufficient to identify the general mechanisms governing deduction, nor are they intended to do so.

A common property of the theories explored above is their limited scope: they are actually microtheories and lack a unified view of the deductive process. However, the search for a unified mechanism for deduction has been the topic of some recent studies. Oaksford and Chater (1994) pursue this perspective yet further. According to their proposal, deductive competence is part of the general thought processes, which are explained by *rational analysis*. The intuitive idea is to apply information theory to the evaluation of various reasoning tasks, and then devise strategies that satisfy these informational accounts (see the model proposed by Chater & Oaksford, 1999 for syllogistic reasoning).

Among the theories which focus on the issue of deductive competence and its general properties, there is the theory of Rips (1994) in the mental logic paradigm. He claims the existence of a central deductive mechanism devoted to formal reasoning; the mental processes involved are reproduced by a system named PSYCOP, which is implemented in Prolog. When presented with premises-conclusion pairs, PSYCOP uses a set of formal rules to construct the lines of a proof. Among the models based on formal rules of inference, this is the best account of both syllogistic and propositional reasoning. However, it has the major limitation of not accounting for consistent patterns of erroneous inferences (see Johnson-Laird, 1997). As far as developmental issues are concerned, they have not been considered by Rips, but extensions are possible.

The Verbal Reasoning Model, proposed by Polk and Newell (1995) within the Soar unified architecture (Newell, 1990), deserves special consideration. They argue that syllogistic reasoning relies on the encoding and decoding of the verbal information given in the premises. More generally, the authors claim that there are no specific mental processes devoted to reasoning: language comprehension and generation can give an account of the entire range of deductive phenomena. Their proposal is an alternative theory within the model-based paradigm (see below), but it states that the search for counterexamples plays little or no role in syllogistic reasoning (however, for a counter argument see Bucciarelli & Johnson-Laird, 1999). The scalability of Soar is well demonstrated, and in principle it might be extended to incorporate developmental aspects of syllogistic reasoning.

Finally, models of reasoning complying with the Mental Model Theory have been proposed for syllogistic, propositional and relational reasoning separately; all of them are supported by experimental evidence, and the latest contributions also give a developmental account of the subjects' performance (Bara, Bucciarelli & Johnson-Laird, 1995; Bara, Bucciarelli, Johnson-Laird & Lombardo, 1994; Johnson-Laird, Oakhill & Bull, 1986). Moreover, MMT has been successfully applied to probabilistic reasoning (Johnson-Laird, Legrenzi, Girotto, Sonino-Legrenzi & Caverni, 1999; Johnson-Laird & Savary, 1996), temporal reasoning (Schaeken, Johnson-Laird & d'Ydewalle, 1996; Vandierendonck, De Vooght & Dierckx, 2000), causal reasoning (Geminiani, Carassa & Bara, 1996; van der Henst, 1999), modal reasoning (Bell & Johnson-Laird, 1998; Goldvarg & Johnson-Laird, 2000), counterfactual thinking (Byrne, 1997; Byrne & McEleny, 2000), pragmatics (Manktelow, Fairley, Kilparick & Over, 2000; Sperber, Cara & Girotto, 1995), decision making (Devetag, Legrenzi & Warglien, 2000; Legrenzi, Girotto & Johnson-Laird, 1993), along with other sorts of thinking (for a review, see Garcìa-Madruga, Carriedo & Gonzalez-Labra, 2000; Garnham & Oakhill, 1996). This versatility favors MMT in relation to other theories.

In what follows, we provide a unified model of deductive reasoning following the basic tenets of MMT. The model is fully consistent with the microtheories developed within the mental model framework. Its novelty consists in a parsimonious account of deductive ability across deductive tasks and age groups. Our proposal satisfies the three dimensions which assess the acceptability of a theory of deductive reasoning, as defined at the beginning of this section.

3. Ontology of mental models

The aim of this section is to translate into formal terms the intuitive tenets of MMT. Though we are aware that the story is far from being complete, a formal system is necessary for the development of any theory that claims to be computational. Our effort is towards a formalism with a degree of generality that goes beyond the specific features of the model described in this paper. We shall proceed by providing an intuitive description, followed by formal definitions.

A model is a specific mental representation whose structure is analogical with respect to the state of affairs it represents. There is a direct correspondence between the entities and the relationships present in the representational structure (the model) and the entities and the relationships of what is being represented (the state of affairs in the real world).

The construction of a model strongly depends on its intended use. So, several models may refer to a single state of affairs, as several tasks may be required to represent that state of affairs: reasoning, planning, discourse comprehension, and so forth. An assertion may be usefully translated into several models, even though it is presumed that only one of these models provides the best description of the state of affairs, given the system's goals. Each model highlights those aspects of the state of affairs that are fundamental for the task at hand; in this sense, the result is to be underspecified with respect to the actual state of affairs (for the notion of *underspecification,* see van Deemter & Peters, 1996). A mental model depicts what is useful or interesting for the system to stress at a particular moment. No unique optimal representation exists of any entity. What does exist is a representation geared to the system's goals: mental models exhibit values for use, not absolute values.

The underspecification of a mental model renders it compatible, in terms of structure and content, with several states of affairs in the real world. We express this feature by saying that a mental model represents a *set of states of affairs.* This set is formed by all the states of affairs that have in common both the tokens and the relations of the model. However, when a model is utilized in a specific situation, as a mental representation useful for performing a task, it assumes a single interpretation, because it is meant to refer to a specific state of affairs in the real world.

To sum up, a *model* represents a set of states of affairs which concerns entities of the world; a model expresses what the states of affairs in the set have in common in terms of content and structure. The *tokens* are the participants to the relations that describe the states of affairs. In turn, tokens can be either *elements* (that directly refer to mental representations of entities of the world generated from internal representations or external inputs), or *encapsulated models* (that refer to states of affairs considered as a whole). Therefore, the definitions of token and model are circular, since they mutually refer to each other: the sense of such a circularity lies in the fact that a model can be considered either as a decomposable entity, or as an atomic entity within another model.

Mental Model Theory claims that reasoning consists in constructing and manipulating mental models, formulating putative conclusions supported by those models, and validating such conclusions (see e.g. Johnson-Laird & Byrne, 1991). We propose a more analytic version of these stages: Construction, Integration, Conclusion, Falsification and Response (Fig. 1). The integration procedures and part of the falsification procedures are task independent; they form a core system which we call Millstone (see below). The term is borrowed – roughly with the same meaning – from Babbage's analytic engine.

3.1. Construction phase

In this phase linguistic and perceptual inputs are translated into analogical representations. Let us proceed with an example for each of the domains we are considering. Because our aim is to propose an all-purpose ontology, the notation presents some differences with respect to the ones historically introduced in the mental model literature from the 1983 up to now.

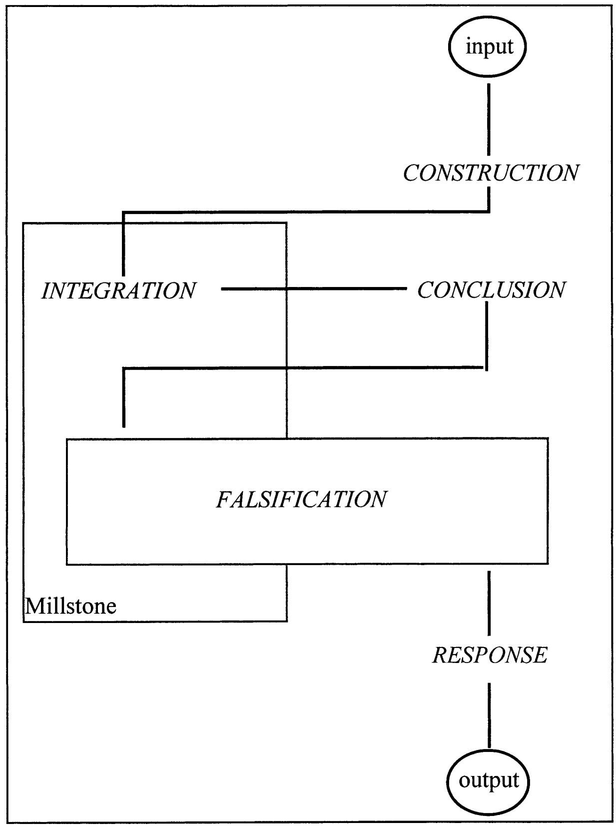
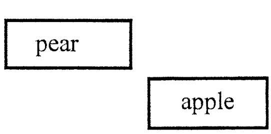


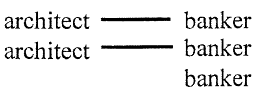
Fig. 1. The five phases of deduction: Construction, Integration, Conclusion, Falsification and Response.

In propositional reasoning, a premise containing the *exclusive or* connective, such as 'Either there is a pear or an apple, but not both', can be represented by constructing two models of the following kind:



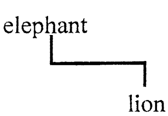
Each model contains a single token (strings of characters) and no relations. We introduce the conventional notation of inscribing each model in a box, except when we deal with a single model. Models are ordered left to right according to the sequential order in which the states of affairs they represent are described in the sentence. Models, and tokens inside models, are also conventionally arranged top-down. Note that the two models lie on different lines to enhance the fact that they never occur together.

In syllogistic reasoning, a possible analogical representation of the quantified premise ‘All of the architects are bankers’ is the following single model:



There are five tokens (strings of letters) and two relations (segments). Each line denotes a separate individual. Note that the representation of the premise includes bankers who are not architects.

In relational reasoning, a possible analogical representation of the premise ‘The elephant is taller than the lion’, referring to a comparison between two entities with respect to some dimension, is the following single model:

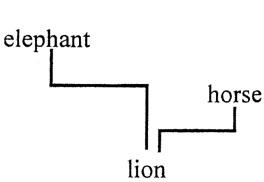


The relation involving the two tokens is represented by the broken line; the respective height of the two animals is analogically rendered by the ‘elephant’ standing above the ‘lion’. The same representation can account for a number of comparisons (taller, faster, larger, . . .), when the task is relational reasoning.

From these considerations on the analogical structure of the model representations, it is clear that 1) a model must include a grid of positions that are assigned to tokens (a *matrix*); 2) tokens must have a name; 3) some tokens may be in relation. A system based on model representations must include the abilities to generate tokens and to establish relations between them.

3.2. Integration phase

In this phase all the models of the premises are integrated into a single mental model (integrated model) by overlapping their identical tokens. For instance, if we consider the model of the premise ‘The elephant is taller than the lion’ together with the model built from the premise ‘The horse is taller than the lion’, a possible integration of the two results in the following model:



Thus, integration is accomplished by merging the models of the premises to yield a new model which presents as a kernel one token (lion) that results from the *Match* of two tokens, one per model. The integration procedures are task-independent.

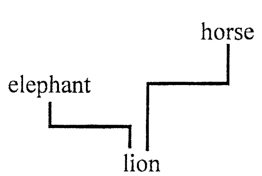
3.3. Conclusion phase

In this phase reasoners take into account the specific deductive task at hand. They extract the relevant information from the model produced in the integration phase, and formulate a putative conclusion. The type of conclusion can be different depending on the task: it can be an inference, a truth-value judgment ('true' or 'false'), or an action to be performed. In the example above, the integrated model supports the conclusion 'The elephant is taller than the horse'. To formulate a conclusion it is necessary to transform the requirements of the specific task into cues for *extracting* the relevant information from the integrated models.

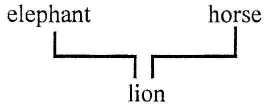
3.4. Falsification phase

In this phase, which MMT considers the core of human rationality, the reasoners attempt to *falsify* the putative conclusion previously obtained, by searching for alternative integrated models which are *inconsistent* with the conclusion. If the search fails, then the conclusion is valid; if the search is successful, the reasoner formulates a new conclusion which is consistent with all the models produced. When it is impossible to formulate a conclusion, the reasoner asserts: 'No valid conclusion'.

MMT claims that people are rational in principle, but fallible in practice. People are rational because they grasp the need to falsify a conclusion; a conclusion is valid only if there is no way in which it can be false given that the premises are true. People are fallible since the actual application of the falsification procedure may fail. In the example above, there exists an alternative possible integration of the models of the premises, namely:



Such a model supports the conclusion 'The horse is taller than the elephant'. Moreover, it is also possible to construct a third integrated model where the elephant and the horse have the same height, which supports the conclusion 'The elephant is as tall as the horse':



As none of the conclusions is consistent with the three possible integrated models of the relational premises, the correct answer is ‘No valid conclusion’. Thus, either people exhaust the set of the possible integrated models of the premises in the example, check that no conclusion is true in all of them, and conclude ‘No valid conclusion’, or they stop at one of the integrated models, and draw an erroneous conclusion.

Falsification consists of two procedures: the first one *searches* for integrated models that are alternative to the first one produced in the integration phase. The second procedure checks both the *consistency* of a conclusion with a model, and the *equivalence* between two models. The search for alternative models is task-independent; the tests of Consistency and Equivalence are task-dependent. The test of Consistency discards a conclusion which is inconsistent with an integrated model, and attempts to formulate a new conclusion which is consistent with all the integrated models; in case of failure, it returns ‘No valid conclusion’. The goal of the test of Equivalence is to discard the models which are equivalent (with respect to the task) to one previously generated, because they are redundant.

3.5. Response phase

In this phase, the reasoners translate the conclusions from the internal model format into a communicative behavior. This phase is the converse one of Construction, where the Interpreter translates the premises into the model format. Operatively, the reasoner *generates* the (linguistic or motor) response that expresses the conclusion, by taking into account the task.

Let us move to the formal notation. Mental model representations feature a mixture of linguistic and graphical forms.[6]

A model is a triple <T, R, A>, where T is a nonempty matrix of tokens, R is a set of relations on T, and A is a set of annotations. Tokens are ordered left-to-right in the matrix, according to the input order. This left-to-right order is meant to reflect an operative feature of the working memory. The columns of the matrix are ordered left-to-right and the tokens in the same input position are located in the same column. When a premise requires several models, the various matrices are stacked one above the other; the tokens respect the input order through the models (for an example, see the models introduced above for the premise ‘Either there is a pear or an apple, but not both).

An **element** is a pair <S, A>, where S is a symbol taken from a given vocabulary, and A is a set of annotations. In the scope of our work, the symbols are words of natural language, that is, strings of letters (architect, banker, elephant, . . .).

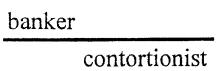
A **token** is either a model or an element. It is uniquely identified by its coordinates in a matrix.

A **relation** is an ordered sequence <r, x_1, . . . , x_n, A>, where r is a relation symbol (taken from a set of symbols), x_1, . . . , x_n are tokens in T, A is a set of annotations. Here are examples of relevant relations:

- ‘connected with’ (**CW**). CW is a relation that forms an individual by ‘connecting’ two of its properties. For example, <CW, architect, banker, ϕ>, represented graphically as architect —— banker

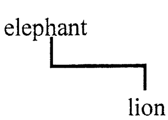
denotes an individual which is both an architect and a banker. ϕ means that the set of annotations is empty.

- ‘never connects with’ (**NCW**). NCW is the relation that states that two properties cannot hold for the same individual. For example, <NCW, banker, contortionist, ϕ>, represented graphically as



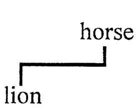
denotes two different individuals, one banker and one contortionist. The model states that an individual who has the two properties cannot exist.

- ‘above’ (**AB**). AB is the relation that states that one token must occupy in the matrix a structural position higher than that of another token. For example <AB, elephant, lion, ϕ>, represented graphically as



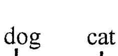
constrains the position of ‘elephant’ to be higher than the position of ‘lion’.

- ‘below’ (**BE**). BE is the relation that states that one token must always occupy in the matrix a structural position lower than that of another token. For example <BE, lion, horse, ϕ>, represented graphically as



constrains the position of ‘lion’ to be lower than the position of ‘horse’.

- ‘same height’ (**SA**). SA is the relation that states that two tokens must always occupy in the matrix a structural position at the same height. For example <SA, dog, cat, ϕ>, represented graphically as



constrains the positions of ‘dog’ and ‘cat’ to be at the same height.

An **annotation** is a symbol of the set **{. . . , P, not}**. An entity (model, element, relation) with a nonempty set of annotations is said to be *annotated.* Annotations are propositional enrichments of the analogical structure of a mental model; for a defense of annotations within models, see Polk and Newell (1995). Each annotation has a different meaning when used with a specific entity, but not all the combinations are possible.

1. ‘. . . ’ can only annotate models. ‘. . . ’ specifies that a model contains some implicit information (tokens and relations). A model <T, R, A> such that ‘. . . ’ $\in$ A is said *implicit;* a model <T, R, A> such that ‘. . . ’ $\notin$ A is said *explicit.* ‘. . . ’ does not affect the truth conditions of a model, which remain exclusively based on the overt information.
2. **P** expresses the potentiality for some representational entity (element or relation). Potentiality is an indeterminacy with respect to the entity itself. A *potential* element or relation is an entity whose existence in a model is uncertain. For example, if we annotate the relation CW introduced above with P, we yield a relation that states that two properties can potentially hold for a single individual. For example <CW, architect, banker, P>, represented graphically as

   architect ----- banker

   denotes a potential individual which is both an architect and a banker. Actually, it could be either a single individual (both architect and banker) or two distinct individuals. To formulate a conclusion based on a model with a CW relation annotated with P, the reasoner works on two models, one with a CW relation and one without (cf. Johnson-Laird & Bara, 1984, on annotated tokens).
3. **not** applies to any entity. When not annotates a model, this means that the set of states of affairs represented by the model is not the case. The scope of such negation is the model which it annotates, that is, the set of tokens and relations that compose the model. When **not** annotates an element, this means that the element is absent in a model. Finally, when **not** annotates a relation, this means that the relation is not the case.

The following property holds: a model cannot represent contradictory states of affairs. As a consequence, the tokens t and not-t cannot appear in the same model. Such a property is a fundamental assumption in our theory. Contradictions are definitely possible within knowledge representation. People often hold contradictory beliefs without knowing it - and sometimes even knowing it (Elio & Pelletier, 1997). But complications arise as soon as one becomes aware of two contradictory states of affairs. E.g. in perception, Gestalt psychologists showed the impossibility of simultaneously perceiving two figures in figure-ground organization (Rubin, 1921); in thinking, Braine and Rumain (1983) showed that people judge as absurd the co-occurrence of a state of affairs (p) and the fact that the same state of affairs is not the case (not-p). It is virtually impossible to maintain an overt contradiction. Our formalization of MMT represents such a natural constraint through the property above; that is, it allows the possibility of contradictory states of affairs only when represented by different models. In fact, we implement falsification through the capability of detecting contradictions among models (see below).

1. to generate a token that represents an entity either perceived or described, or produced from memory.

2. to establish relations among tokens in order to represent a state of affairs.

3. to expand a token into a model.

4. to shrink a model into a token.

5. to negate an entity (token, relation) in order to represent the fact that it is not the case.

6. to flesh out an implicit token (Flesh-out token).

7. to flesh out an implicit model (Flesh-out model).

8. to detect identities/differences among tokens (Identity/Difference).

9. to identify referents of descriptions (Referent).

10. to alter the order in the sequence of the models of the premises (Re-order).

11. to alter the order of tokens in a mental model (Invert).

12. to change the position of some movable tokens (Token-moving).

Fig. 2. The basic abilities involved in deductive reasoning.

3.6. Basic abilities

The basic abilities for constructing and manipulating tokens are listed in Fig. 2. We follow the criteria given by Bara et al. (1995) for establishing what can be considered a basic operation in constructing and manipulating a mental model. They assume that such operations are innate predispositions whose emergence is caused by the increase of both cognitive resources and experience. From a compositional point of view (see Sternberg, 1985), the basic operations are elementary processes that operate upon representations.

Bara et al. (1995) provide experimental evidence for some of the basic operations (see below).

Let us review the abilities: the numbering corresponds to Fig. 2. The abilities 1–5 have been postulated on the basis of shared intuition among model theorists. They are the essence of the capability to represent the world according to MMT.

3.6.1. Abilities 1 and 2

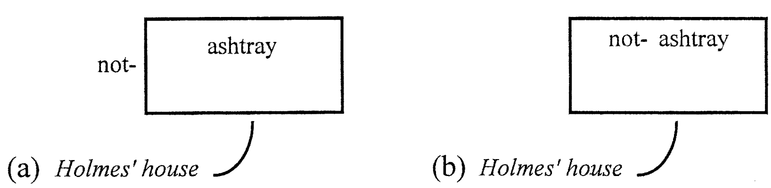
There are three possibilities for construction of elements and relations. An element or a relation may refer to a mental representation constructed from external inputs, by perception (the vase presently in front of me) or by verbal description (the car a friend is talking about). Alternatively, they are constructed from pre-existing internal representations of entities, that is, by memory (my first watch).

3.6.2. Abilities 3 and 4

These abilities derive from the circular definition of model and token. As a model in turn can be an element or an encapsulated model, such abilities permit the transition from the former to the latter, and vice versa. Ability 3 expands a token into a model, to obtain a decomposable entity composed of tokens and relations. Ability 4 shrinks a model into a token, to obtain an encapsulated model that may be handled as a whole.

3.6.3. Ability 5

A negative representational entity is a token (element or model) or a relation annotated with 'not'. A negative element <E, not> in a model expresses the fact that the state of affairs represented by the model does not include an element E. A negative model expresses a state of affairs which is not the case. In constructing a mental model, the reasoner may generate a negative token after detecting the absence of an entity in the perceived or described state of affairs (e.g., through overt mention). For instance, the model built from the assertion 'In Holmes' house, there is not an ashtray' can be one of the following two, where the choice between the two depends on the specific reasoning task:

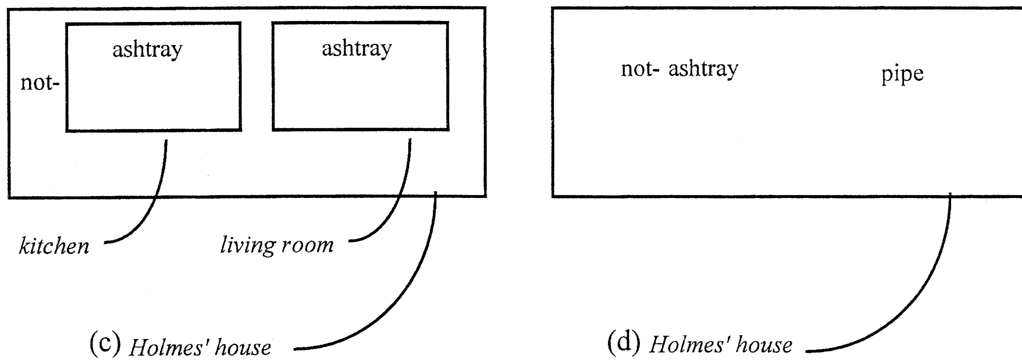


(a) is the model representation of the state of affairs 'In Holmes' house there is an ashtray'. The model is annotated with *not,* to claim that the state of affairs represented by the model is not the case. (b) is the model of the state of affairs 'In Holmes' house there is not an ashtray'. In this case, it is the element 'ashtray' to be annotated with *not,* in order to negate its presence within the model of Holmes' house. The model represented in (b) is more sensible if the intention of the reasoner is to mark the contrast between objects which are and objects which are not in the house. As a further clarification of the scope of negation, consider now the following models:

In (c), which is a model of the assertion 'In Holmes' house there is not an ashtray in the kitchen; there is one in the living room', the scope of the annotation 'not' is the encapsulated model representing the state of affairs 'In the kitchen of Holmes' house there is an ashtray'. In (d), which is a model of 'In Holmes' house, there is not an ashtray, but there is a pipe', the scope of the negation is the element 'ashtray' in the model of Holmes' house.

A negative relation expresses the fact that the relation itself does not hold between two tokens. For example, to represent the assertion 'There is an architect who is not a banker', we can use the negative relation <CW, architect, banker, {not}>.

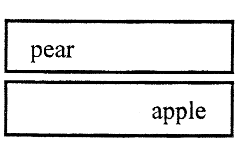
Although our computer model is not concerned with the Construction phase, we would emphasize a fundamental assumption about the mental processes underlying the representation of negation: the negation of a state of affairs is more difficult to represent than the state



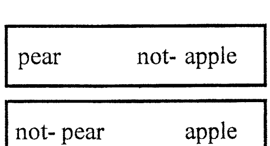
of affairs itself because, in the former case, people have first to construct the representational entity to be negated, and then to negate it. Indeed, Mental Model Theory postulates the principle of truth as guiding the construction of models. According to this principle individuals have no direct way to represent what is false, but derive the false cases from the true ones (see the experimental evidence in Barres & Johnson-Laird, 1997, and Sacco, Bucciarelli & Adenzato, 2001).

3.6.4. Ability 6

In the manipulation phase, it is possible to introduce a token when the representation contains that token implicitly. Some information is implicit when it is in principle accessible to the reasoner, since it corresponds to established knowledge, but it is not currently available in working memory. E.g., consider the information that is implicit in the two models for a disjunctive assertion like ‘There is either a pear or an apple, but not both’:



The knowledge of the meaning of the disjunction allows the reasoners to make the model representation fully explicit, by introducing the negative tokens (elements) which were implicit:



3.6.5. Ability 7

The ‘. . . ’ annotation in a model means that the premise describes several states of affairs (models), of which only one is currently explicit. The further models can be fleshed-out later on. This occurs, for instance, in dealing with the conditional connective. The initial representation of the assertion 'If there is a circle on the blackboard, then there is a triangle' contains some implicit information, represented by the three dots annotation (Johnson-Laird & Byrne, 1991):

circle	triangle

...

The representation indicates that the assertion is true when both the circle and the triangle are on the blackboard. Additionally, the three dots indicate that other states of affairs where the assertion is true may be recovered from memory. The result is the following explicit representation:

circle	triangle
not- circle	not- triangle
not- circle	triangle

According to such a representation, the assertion 'If there is a circle on the blackboard, then there is a triangle' is true even when neither the circle nor the triangle are on the blackboard (2nd model), and when the triangle is on the blackboard, but the circle is not (3rd model). In order to be able to flesh out implicit models, reasoners must know the meaning of the connectives.

3.6.6. Abilities 8, 9, 10, 11, 12

All these abilities are analyzed in detail in the following section, where we deal with the reasoning procedures. Here, we just mention two sorts of experimental evidence that support the psychological plausibility of these abilities. First, the ability to detect identities and differences strongly accounts for the capacity to draw both syllogistic inferences (Bara et al., 1995), and relational and propositional inferences (Bara, Bucciarelli & Lombardo, 2000). Second, Reorder and Invert operations affect syllogisms' difficulty (Bara et al., 1995).

4. A computational theory of reasoning based on mental models

In this section we present the global theory, and UNICORE (UNIfied COmputational REasoner), the computational program devised from it. Then, in section 5, we deal with the resources that constrain the run of the system, and that can explain erroneous performance

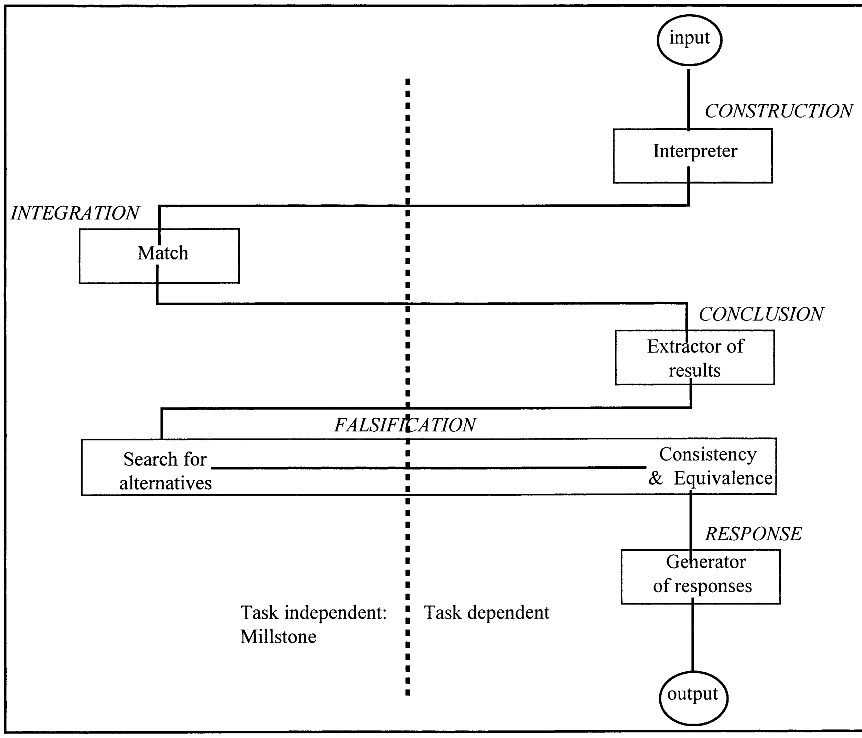


Fig. 3. The reasoning process. The procedures of the Millstone (Integration and Search-for-alternatives) are invariant through deductive domains. The procedures outside the Millstone (Construction, Conclusion, Consistency-&-Equivalence and Response) are task dependent.

in reasoning. Sections 6–8 are devoted to comparing the performance of human subjects versus artificial ones, in syllogistic, relational, and propositional reasoning, respectively.

The general schema of the reasoning process, which lists the flow of control through the main components of UNICORE, is illustrated in Fig. 3. The reasoning process can be analyzed at different levels of granularity. At an abstract level, MMT provides a general framework in which the core of the reasoning process is performed by interpreting premises, integrating models, formulating putative conclusions, falsifying them, and generating a response. Construction, Integration, Conclusion, Falsification and Response are high-level procedures that operate serially.

On a fine grain, the basic abilities of Fig. 2 provide the low-level procedures that manipulate the tokens and the relations inside the models. The implementation of the high-level framework through these low-level procedures is accomplished via specific strategies. A strategy is a specific realization of the general processes described by MMT: it dictates the application order of the low-level procedures. UNICORE implements a specific strategy which is in accordance with the general tenets of the theory; as a matter of fact, MMT allows for a variety of possible strategies (Byrne, Handley & Johnson-Laird, 1995; Johnson-Laird, Savary & Bucciarelli, 2000). In Fig. 4 we illustrate the procedures (gray frames), together with the data they process (black frames). We shall adopt this convention for all the figures.

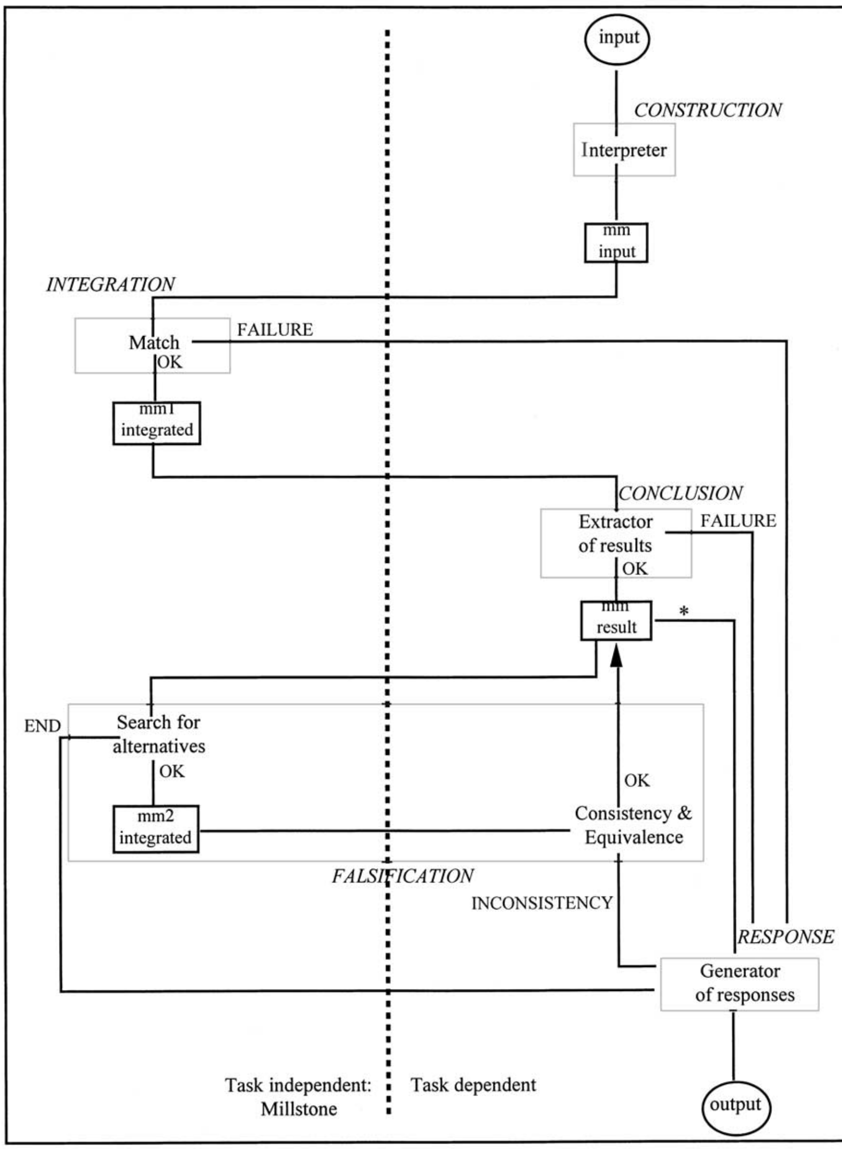


Fig. 4. The computational model of the reasoning process: UNICORE (UNIfied COmputational REasoner). Procedures are within gray frames, and data are within black frames. * marks premature exits from the reasoning process.

The *Interpreter* implements the Construction phase. Interpreter stands for the set of capabilities that process both the linguistic premises (compositional semantics) and the information coming from perception or restructuring of internal knowledge. It produces as output a model representation of the input (mm_{input}). In this paper, the Interpreter is considered as a black box that translates the premises into their propositional representations, and then the latter into the mental models that are passed to the reasoning mechanism. The construction of models is outside the scope of the paper. We shall analyze neither the level of neural structure that provides symbols from perception, nor the translation of linguistic expressions into models. As a consequence, UNICORE does not take into account possible incomplete representations of the input, due to the lack of general or linguistic knowledge.

Match implements the Integration phase: it takes as input the mental models provided by the Interpreter, and tries to return an integrated model ($mm1_{integrated}$). An integrated model contains information that was implicit in the input models. In case of failure, the control passes directly to the Response phase, which acts consequently (see below).

The *Extractor-of-results* implements the Conclusion phase. The Extractor selects the items of an integrated model that are relevant to the task at hand. It takes as input an integrated model and produces as output a model which contains only the tokens and relations which represent a putative conclusion, given a task (mm_{result}). For example, in the case of syllogistic reasoning, it marks the tokens and relations concerning the end terms. This result model becomes the current model in the working memory. The asterisk in the figure (*) indicates that it is premature to exit at this point, because the conclusion could be wrong: in fact, the conclusion which is based on the first integrated model might be falsified by a further integrated model of the premises. In case of failure, the control passes directly to the Response phase, which acts consequently (see below).

Search-for-alternatives and *Consistency-&-Equivalence* implement the Falsification phase. This phase takes as input a putative conclusion (first result model), and gives as output a validated result model. Each time it is invoked, Search-for-alternatives tries to produce a new integrated model (if any) of the premises. This integrated model is then passed to Consistency-&-Equivalence. Consistency-&-Equivalence first invokes the Extractor, which selects the relevant information depending on the task, and produces a new result model; then Consistency-&-Equivalence detects the possible presence of contradictions between the current result model in the working memory and the new result model just yielded. If it does not detect any contradiction, but it realizes that the new model supports the same conclusion as the current result model, Consistency-&-Equivalence leaves the working memory inaltered. Otherwise, if the new model does support a looser conclusion than the current result model, it replaces the current result model with the new result model (see below for the definition of ‘looseness’). Finally, if it detects a contradiction, that is, if the two models support inconsistent conclusions, it generates the flag INCONSISTENCY and passes it to the Response phase. When Search-for-alternatives does not produce any more integrated models (flag END), the control goes to the Generator-of-responses.

The *Generator-of-responses* implements the Response phase. In case of success of the previous phases (flag END), it takes as input the current result model in the working memory and translates it into linguistic or motor responses. Otherwise (either the flag FAILURE or INCONSISTENCY), it interprets the failure according to the task.

Before describing these procedures in detail, we analyze the reasoning schema from the perspective of a unified computational model. A unified computational model includes a core of procedures that are invariant across deductive tasks. The invariant procedures of UNICORE are Match and Search-for-alternatives (left-hand side of Fig. 3 and 4). Match returns a first integrated model; Search-for-alternatives returns further integrated models, as required by the reasoning process. We have grouped the two procedures in a unique engine, called

Millstone. The internal processes of the Millstone have no access to the task at hand; they implement the ability to navigate the space of the integrated models of the premises, which is determined by the initial models of the premises and the basic abilities that manipulate models (cf. Polk & Newell, 1995).

The Interpreter, the Extractor-of-results, the Consistency-&-Equivalence, and the Generator-of-responses are task dependent (right-hand side of Fig. 3 and 4). Interpreter and Generator are the communicative interfaces of the reasoning system: in input, a premise can be represented in many ways (Interpreter); in output, a model can support many responses (Generator). The communicative interfaces (Interpreter and Generator), the selection of relevant tokens and relations (Extractor), the checks of consistency and of equivalence (Consistency-&-Equivalence) can only occur with respect to a specific task.

A technical remark: UNICORE is implemented with a modular architecture in C language, and it does not require specific platforms. The architecture accurately reflects the schemas in the paper; conversely, all the presented predictions result from executions of the program, with no hidden parameters. We come back on these issues in section 5. In the rest of this section we describe the various procedures in detail, by referring to the schema in Fig. 4, which illustrates both the flow of control and the changes in the data. The algorithms in pseudocode can be found at "http://www.elsevier.nl/gej-ng/10/15/15/show/."

4.1. Construction

The first phase of the reasoning process is Construction. It takes as input the linguistic or perceptual premises of the task at hand and returns as output their model representations. Given syllogistic, relational or propositional premises, UNICORE assumes that the reasoner constructs suitable representations: they are introduced in the sections devoted respectively to syllogistic, relational and propositional reasoning (i.e., sections 6, 7 and 8).

4.2. Integration

The Integration procedure takes as input the models of the premises and returns as output an integrated model. This integrated model, which is the first result of the exploration in the space of the integrated models, is then passed to Conclusion, in order to extract the relevant information for a specific deductive task (e.g., syllogistic inferences, relational inferences, propositional inferences). If Integration fails to produce any integrated model, the reasoning process yields a failure, which interrupts the reasoning process. In Fig. 5 we illustrate the specific strategy that implements Integration.

Integration invokes three lower level functions: Match, Rearrange and Flesh-out. The input models are juxtaposed one to another, in order to Match the adjacent tokens in different models (rightmost column of the first model, leftmost column of the second model). When Match succeeds, it produces an integrated model. The integrated model is assembled from the input models, first by overlapping the matched tokens and the relations that share those tokens, and then by adding the remaining tokens and relations from the two input models. When Match does not succeed, Integration rearranges the positions of tokens through a set of constrained movements, with the goal of making adjacent some matchable tokens present

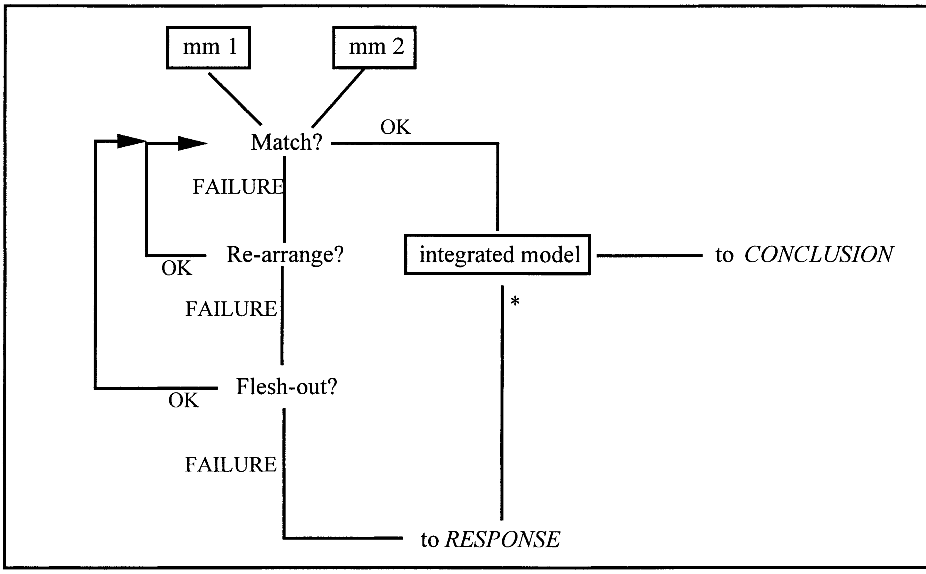


Fig. 5. The Integration procedure, that consists of three functions: Match, Rearrange, and Flesh-out. * marks premature exits from the reasoning process.

in the models. After all these attempts, if some tokens have matched, Integration builds an integrated model; otherwise, Integration tries to flesh-out some implicit information possibly contained in the models of the premises. If Flesh-out succeeds, Integration calls the Match+Rearrange loop again: if Flesh-out does not succeed, Integration definitely fails.

In the following, we describe Match, Rearrange and Flesh-out in detail.

4.2.1. Match

The function Match overlaps the tokens which are identical in the two models of the premises. It takes as input two adjacent tokens t_1 and t_2 in the input models. If the two tokens are identical (ability 8 in Fig. 2: Identity), then Match succeeds, and the resulting token forms the core of the integrated model; otherwise Match fails. In the case of elements, the words that represent the tokens must be identical; in the case of models, all the three components of the triple $<T, R, A>$, must be identical.

In Fig. 6, the function Match is applied to the reasoning domains we deal with in this paper. The tokens that are going to match are in *italics;* the matched tokens, that is, the core of the integrated model, are in **boldface.**

4.2.2. Rearrange and Flesh-out

The function Rearrange manipulates the arrangement of tokens in the models of the premises, in order to give a new chance for the Match function to succeed; thus, it is always invoked after Match has failed. Rearrange consists of a small set of low-level procedures: Reorder, Inversion of the second premise, and Inversion of the first premise. The function Flesh-out makes explicit the implicit information, by invoking two low-level procedures: Flesh-out token and Flesh-out model.

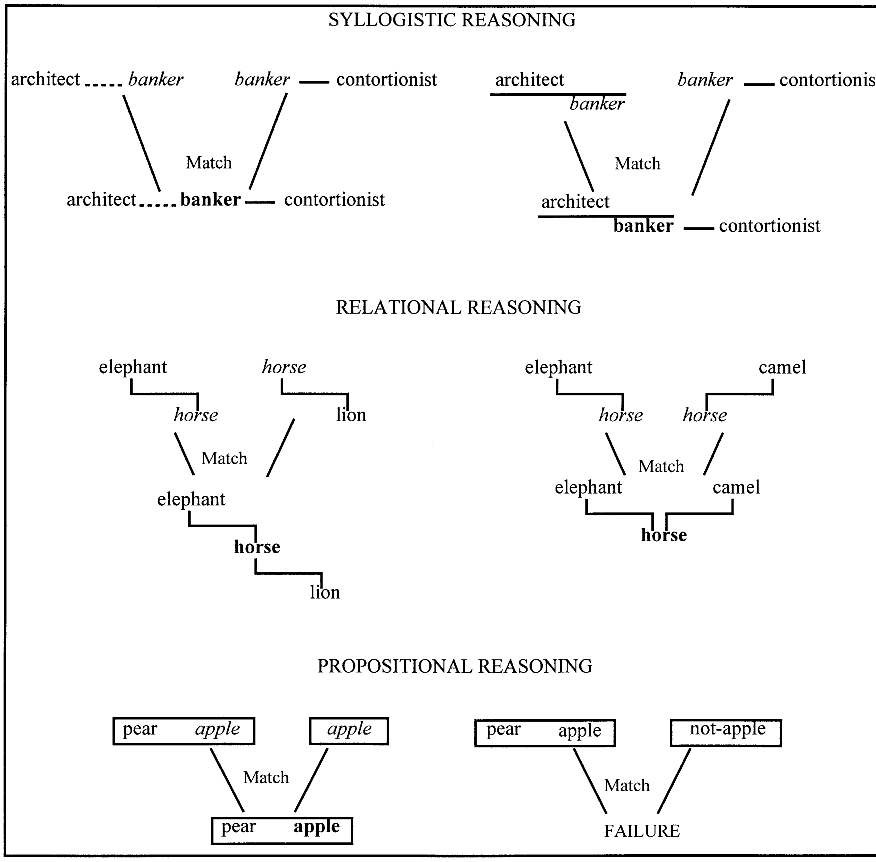


Fig. 6. The function Match in three types of reasoning: matching tokens are in *italics,* the result of Match is in **boldface.** In the last example, Match fails because it does not find any identical tokens.

Reorder (Fig. 7a) inverts the order of the models of the premises.

Inversion of the second premise (Fig. 7b) alters the sequential order of the tokens in the model representation of the second premise.

Inversion of the first premise alters the sequential order of tokens in the model representation of the first premise.

Flesh-out token (Fig. 7c) makes explicit some tokens that were initially absent in the representation, because they had not been triggered by the content of the premises. Consider, for instance, the representation of the assertion 'There is either a pear or an apple on the table, but not both' introduced in the ontology, together with the model of the assertion 'There is not an apple'.

Flesh-out model (Fig. 7d) makes explicit some models which the reasoner had initially kept implicit because of a parsimonious representation of the input premises.

Consider, for instance, the representation of the assertion 'If there is a circle on the blackboard, then there is a triangle' introduced in the ontology, together with the model of the assertion 'There is not a triangle'.[7]

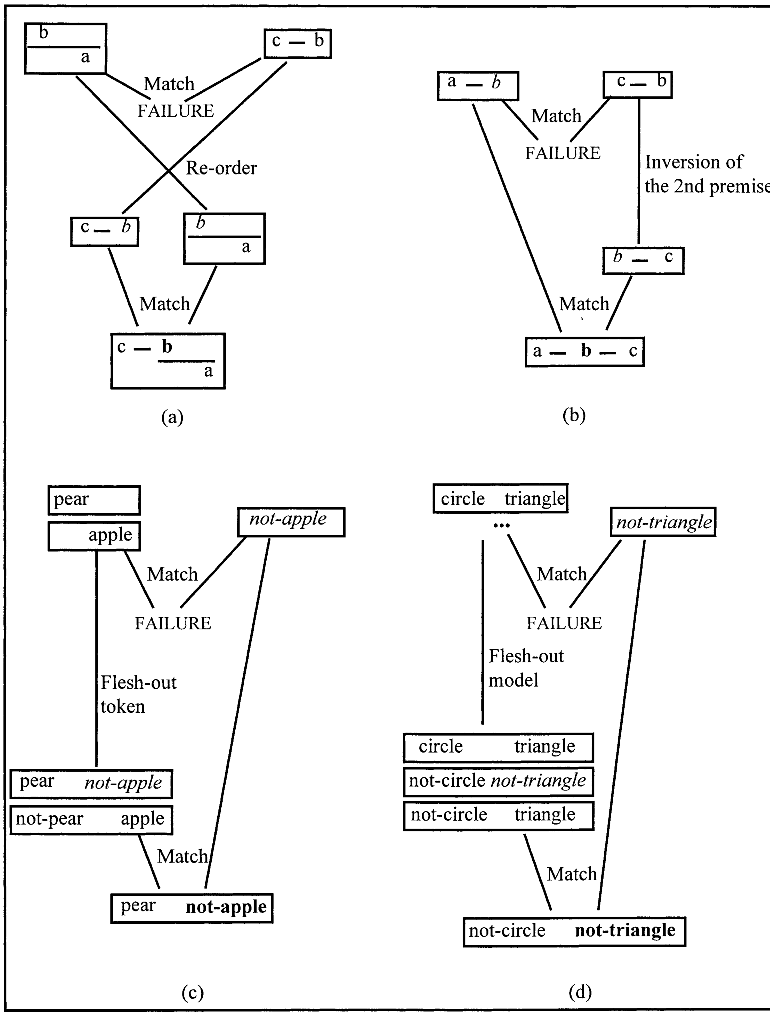


Fig. 7. The functions Rearrange - (a) and (b) - and Flesh-out - (c) and (d). Matching tokens are in *italics*, the result of Match is in boldface.

The degree of difficulty of the functions Rearrange and Flesh-out depends on the amount of information that is manipulated, and on the organization of working memory space. In fact, according to principles of cognitive economy - which dictate to represent the minimal amount of information required to accomplish a task - the reasoner is supposed not to immediately access the whole body of relevant information. Hence, we can postulate the following relative difficulties:

- Rearrange is easier than Flesh-out, because the latter involves the retrieval of information from long term memory.
- Within Rearrange, Reorder is easier than Inversion of the second premise, which is easier than Inversion of the first premise. The rationale is that Reorder applies to

chunks in working memory, whereas Inversion operations apply to the items forming a chunk. Further, the most recent model in memory (second premise) is the most accessible. Thus, inversion of the 2nd premise is easier than inversion of the 1st premise.

- Within Flesh-out, Flesh-out token is easier than Flesh-out model, since tokens are simpler entities than models to deal with.

The difference in difficulty among the Rearrange procedures has been shown by Johnson-Laird and Bara (1984). The authors found that the arrangement of tokens inside syllogistic premises is predictive of the difference in difficulty among all the 64 classical syllogisms. As far as flesh-out procedures are concerned, the degree of difficulty that we are assuming is accounted for by many studies on propositional reasoning (see, e.g., Johnson-Laird, Byrne & Schaeken, 1992).

4.3. Conclusion

The Extractor-of-results procedure implements the Conclusion phase. It takes as input an integrated model and it produces a result model, which represents a first putative conclusion. This procedure is task-dependent: it selects the relevant information in a model which could be read out in several ways. It is the task that determines the correct way of reading a model. In case the Extractor cannot select any information which is relevant for the task, it exits with a flag FAILURE. The Extractor is invoked after every successful execution of the Millstone procedures: it considers one integrated model at a time. An integrated model contains new information with respect to the input premises; the Extractor views this new information in the context of the specific task, and *extracts* it to yield a *result model.* The reasoning program maintains only one result model in the working memory, the *current* result model. The result model which is current after the exhaustion of the space of the integrated models or the failure flag will be passed to the Response phase.

The extractor operates as follows:

1. In syllogistic reasoning, the goal is to relate the end terms (of types a and c), given the indirect links provided by the middle term (of type b) in the integrated models. The Extractor accomplishes this goal by connecting the end terms through the application of transitivity to the relations in the model (see the details of the application of transitivity in the simulation code at "http://www.elsevier.nl/gej-ng/10/15/15/show/"). Ultimately, the result model contains only tokens of types a and c.

2. In relational reasoning (i.e., three-term series problems), the task is similar to syllogistic reasoning, since we have to relate two end terms through the application of transitivity.

3.1. Inference: in this task the reasoner must extract the new information contained in the premise involving a connective, with respect to the situation described by the atomic proposition. The Extractor selects in the integrated model those tokens and relations that are not mentioned in the situation described by the atomic propositions. Such tokens and relations form the result model.[8]

3.2. Truth value judgment: in this task the reasoner must assign a truth value to a specific situation, with respect to a statement. The Extractor verifies that the integrated model

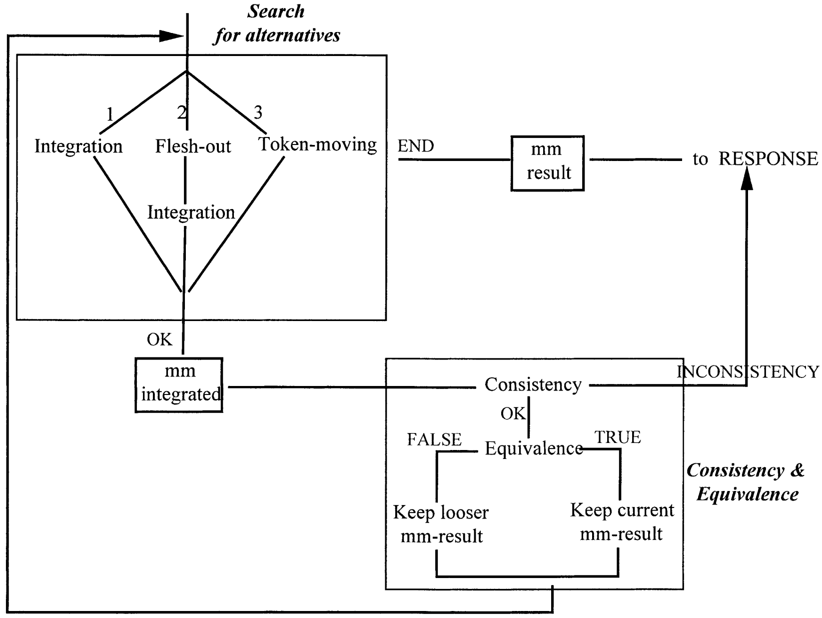


Fig. 8. The Falsification procedure: first, the construction of alternative integrated models is attempted by Search-for-alternatives; then, each putative conclusion is evaluated by Consistency-&-Equivalence.

is identical to the model of the first premise and returns a success flag (OK) or a failure flag (FAILURE).

3.3. Action: in this task the reasoner decides whether or not to execute an action. The Extractor selects in the integrated models a commitment to execute an action or not (viz., a token which represents an action).

When the Extractor operates after Integration, the control can flow in two directions: a correct one that goes to Falsification, and an erroneous one (*), that goes to Response and interrupts the reasoning process. This premature exit of the reasoning process, due either to the limited capacity of the working memory or to a poor degree of mastery of Falsification, can explain several data about subjects' erroneous conclusions. The discussion of these predictions is in section 5 and subsequent ones.

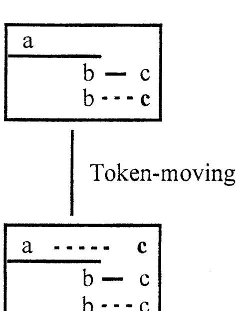
4.4. Falsification

Falsification is accomplished by searching for integrated models of the premises which are alternative to the ones previously obtained, and then checking the validity of each putative conclusion with respect to every integrated model (Fig. 8). Falsification always keeps the current result model (i.e., a conclusion) in the working memory: at the beginning, this is the model which the Extractor produces from the first integrated model. Then, Falsification searches for further integrated models, yields a new result model by extracting the relevant information, and checks whether this new result model is consistent or equivalent with respect to the current result model. If it detects an inconsistency, that is, if the two models represent incompatible states of affairs, then it returns a flag INCONSISTENCY. If two models are consistent, Falsification checks if they are equivalent. If the two models are equivalent, Falsification still keeps the current result model in working memory; if the two models are not equivalent (viz., they are qualitatively different), Falsification determines which is the looser (less constrained) of the two, and keeps it as the current result model. The notions of consistency and equivalence heavily depend on the specific task, and involve the analogical structure: the names and the positions of the tokens, the names of the relations, the number of tokens and relations, and the annotations. Below, we will see what these notions represent in the deductive tasks discussed in the paper. When the search for new integrated models ends, it returns a flag END.

4.4.1. Search-for-alternatives

The goal of Search-for-alternatives is to iteratively yield a new integrated model of the premises that exhibit a different structure, in terms of tokens and relations, compared with the previous ones. Search-for-alternatives has three ways to accomplish this task, which are tried in the following order:

1. to attempt a new integration of the models of the premises, by operating on different tokens with respect to those that have already matched;
2. to flesh out from the models of the premises some implicit information (if any) and then to match again;
3. to modify the analogical structure of the previously integrated models, in case of indeterminacy of some token positions that are not fully constrained by the input premises. Operatively, in this case, Search-for-alternatives duplicates and moves the 'free' tokens in the integrated representations obtained so far (Token-moving). A free token is a token that is not involved in a CW (connected with, Section 3) relation, that permanently ties it to another token. The possible locations of the duplicated token are determined by the constraints posed by the other relations on the analogical structure. A token moved to a new row may potentially be connected with any token on that row: we represent this new state of affairs by overtly introducing a potential CW relation between the token moved and the other tokens of the row. Token-moving provides new integrated representations until no more movements are allowed. In the following example the token c in bold is free:



The Token-moving procedure duplicates the c token and moves it to the first row, thus introducing a potential CW relation with a.

4.4.2. Consistency-&-Equivalence

When Search-for-alternatives returns a new integrated model, Falsification invokes the two-test function Consistency-&-Equivalence, which extracts the relevant information (through the Extractor), and then checks whether this new result model is consistent and non equivalent with respect to the current result model.

If the Consistency test detects an inconsistency, for example, a negative versus an affirmative conclusion, it decides that there is no possible conclusion that is true in both models. In this case, it returns a flag INCONSISTENCY. Otherwise, the two consistent models are evaluated by the Equivalence test. If this test detects a qualitative difference (it returns FALSE), it keeps the looser model in the working memory, and discards the other. Otherwise, the two models are equivalent (TRUE), and the model already in the working memory is maintained. Whichever result model remains eventually in the working memory (when Falsification exits with END), it is the input to Response.

In the following, we define the notions of consistency, equivalence, and looseness for each deductive task.

Consistency:

Two models mm_1 and mm_2 are *consistent* when it does not occur that:

Syllogistic.

1: There are both relations CW (connected with) and NCW (never connects with) on the same pair of tokens.

Relational.

2: There are either both relations AB (above) and BE (below), or AB (above) and SA (same height), or BE (below) and SA (same height), on the same pair of tokens.

Propositional.

3: In mm_1 there is a token x and in mm_2 there is a token not-x.

If one of the conditions 1–3 holds, then mm_1 and mm_2 are *inconsistent.*

Equivalence:

Two models mm_1 and mm_2 are *equivalent* when:

Syllogistic and relational.

The elements in the two models have the same names (even if the two models have a different number of elements), and participate in the same relations (i.e., relations with the same relation symbol and the same argument names), possibly none.

Propositional.

The tokens (models) in the two models have the same analogical structure <T, R, A>.

If two models are *not* equivalent, they are *qualitatively different.*

Looseness:

Given that two models mm_1 and mm_2 are consistent and qualitatively different, the model mm_1 is *looser than* the model mm_2 if mm_1 contains more free tokens than mm_2 in syllogistic and relational reasoning, and if there exists at least a token x in mm_2 which is not in mm_1 in propositional reasoning.

4.5. Response

The Generator-of-responses procedure implements the Response phase. It takes as input the current result model and the flag returned by the previous procedures (FAILURE, INCONSISTENCY, END), and produces a motor or linguistic response, which describes a conclusion, a truth value judgment, or a commitment to an action. The current result model may either have been validated by Falsification, or represent an erroneous result (*) which has bypassed Falsification. The Generator is task-dependent.

If the flag is FAILURE, or INCONSISTENCY, and the task is not Judgment, then the verbal answer is 'No valid conclusion'. In case of Judgment, the answer is 'False'.

In the case of syllogisms, the Generator checks the relations between the tokens and produces one of the following answers:

- 'All x are y', when each token of class x is CW with a token of class y.
- 'Some x are y', when at least one token of class x is CW with a token of class y and at least one x is not.
- 'No x are y', when each token of class x is NCW with each token y.
- 'Some x are not y', when at least one token of class x is NCW with a token of class y, and at least one x is not.
- Otherwise, 'No valid conclusion'.

In the case of relational reasoning, the Generator checks the positions of the tokens and produces as output one of the following answers:

- 'x is more Ω than y', when the position of x on the left in the matrix is higher than the position of the token y on the right.
- 'x is less Ω than y', when the position of x on the left in the matrix is lower than the position of the token y on the right.
- 'x is as Ω as y', when the two positions are equal;
- Otherwise, 'No Valid Conclusion'.

Note that Ω stands for any adjective in the premises (e.g., fast, big, fat, tall, . . .).

In the case of propositional reasoning, the Generator responds differently in the three tasks.

Inference:

- 'There is z', when there is a token z in the model.
- 'There is not z', when the token z is annotated with negation.
- Otherwise, 'No valid conclusion'.

Truth value judgment:

- 'True', when the flag is END.
- 'False', otherwise.

Action:

- ‘Execute action α’, when there is a token in the model which represents the action α.
- ‘Do not execute action α’, when there is a token in the model which negate the action α (not-α).
- Otherwise, ‘No valid conclusion’.

This concludes the description of the unified computational model, where we have maintained a neat separation between the task-independent (Millstone) procedures and the task-dependent procedures. In the sections 6–8 we show how the model operates in the three areas of deductive reasoning we have considered. In the following section we describe the overall constraints that cause the model to produce erroneous outputs.

5. Predictions of erroneous outcomes

Although people possess the complete deductive competence, they may fail. It is a major tenet for a theory to be able to predict systematic errors, as they are observed in the human subjects, without resorting to *ad hoc* assumptions. In this section we shall illustrate how UNICORE predicts the erroneous responses, which correspond to premature exits from the reasoning process (the asterisks in Fig. 4 and 5). In UNICORE there is no parameter that has to be tweaked to get the model to perform as the experimental subjects do. This is a major difference from the other computational models in literature. We have imposed a series of explicit constraints to simulate the performance of subjects of increasing age. However, we have not reproduced individual differences in performance. This would have guaranteed the best possible fit to data, but through *posthoc* adjustments. Data fitting without a theoretical justification of hidden parameters is as easy as it is irrelevant.

Errors occur because of cognitive constraints on each procedure involved in the reasoning process. Although UNICORE may implement constraints at every phase, in the present paper we deal only with the constraints on Falsification, because it is the most significant phase in deduction. They are implemented by UNICORE and give rise to predictions of erroneous responses in all the domains of deduction. Note that these constraints apply only to the problems that require more than one model to be solved; otherwise, Falsification becomes void, because the first conclusion is the correct one.

Falsification relies on a series of abilities and cognitive resources:

1. To grasp the importance of falsification.

The reasoners invoke Falsification because they grasp that the task demands a *necessary*, rather than a *possible* conclusion. Thus, they attempt to falsify the putative conclusion previously obtained.

2. To be able to falsify

The reasoners are able to search for alternative models and to check their consistency and equivalence. They attempt to find a possible integrated model of the premises that is qualitatively different from the ones previously obtained, and to formulate a conclusion which is consistent with all the integrated models.

We shall use the term 'degree of mastery of Falsification' to refer to the individual abilities mentioned in 1 and 2.

3. To have the working memory capacity required by the task at hand.

The reasoner has the working memory capacity necessary to keep in memory all the integrated models, in order to produce a conclusion which is consistent with all of them.

UNICORE implements the prediction that falsification might not occur, due either to a poor degree of mastery (abilities stated at point 1 and 2), or to a poor working memory capacity with respect to the demand of the task (see point 3). The program simulates the behavior of three artificial subjects, who are representative of three different age groups. We hypothesize that children below 10 years do not search for alternative models of the premises because mastery of Falsification has yet to develop. Mastery of Falsification occurs both through mastery of Search-for-alternatives and mastery of Consistency-&-Equivalence.

5.1. Mastery of Search-for-alternatives

After 10 years of age, mastery of Search-for-alternatives takes place in two following steps. At the first developmental step, the reasoners realize the utility of going beyond the first successful integration of the premises, in order to *confirm* the putative conclusion. Hence, the reasoners tend to construct models which are *equivalent* to the first one, thus supporting the conclusion already drawn; it may be considered a confirmation process. The second step goes beyond confirmation, and consists of a search for models which are *qualitatively different* with respect to the previous ones.

To sum up, children start by being totally unable to produce new integrated models of the premises. Then, they construct models which are new, but equivalent to the ones previously obtained. All these models support the initial putative conclusion. Only after adolescence do reasoners search for counterexamples to a conclusion. This assumption is supported by a previous study of Bara et al. (1995), on subjects of analogous age to our adolescents.

The emergence of Search-for-alternatives has to be supported by increasing cognitive resources, and in particular by working memory (wm) capacity. The amount of memory required by deduction affects Falsification, because of the necessity to keep the models in memory while operating on them. If the wm capacity is not sufficient to retain the set of the integrated models of the premises, the conclusion will be based on a subset. Such a conclusion can be erroneous since it is not necessarily consistent with all the models of the premises.

Wm capacity is a good predictor of the performance of subjects belonging to different age groups, since it is well known that it increases with age (Baddeley, 1986). To discretize the increment of wm capacity along age, the program assigns a wm capacity for one model to children, for two models to adolescents, and for multiple models to adults. The rationale is in Bara et al. (1995). They found that over all the participants of different ages who took part in their experiment, the wm capacity accounted for the 8% of the variance in solving deductive problems.

	Artificial subjects		
Constraints	children	adolescents	adults
Mastery of Search-for-alternatives	no	yes	yes
Cognitive Resources	for one model	for two models	for multiple models
Predictions			
The answer is based on ...	the first integrated model	two integrated models	several integrated models
Constraints			
Mastery of Consistency & Equivalence	no	partial	yes
Predictions			
The answer to multiple model problems is...	based on the first integrated model	'No valid conclusion'	based on several integrated models

Fig. 9. Performance errors in UNICORE depend on both the number of models to keep active in the working memory space and the complexity of the manipulation procedures involved. The constraints are in the falsification phase, and they apply only to multiple models problems.

5.2. Constraints in the program

The implementation of these assumptions in UNICORE leads to the simulation of three types of artificial subjects: children, adolescents and adults. The constraints imposed by mastery of Search-for-alternatives and cognitive resources together, support the following predictions about the performance of subjects belonging to different age groups (Fig. 9):

Children (confirmation, and wm capacity for one model) will base their response on the first integrated model of the premises.

Adolescents (Search-for-alternatives, and wm capacity for two models) can in principle construct alternative integrated models of the premises because they master Search-for-alternatives. However, as they have a wm capacity for two models, they will base their response on just two of the integrated models possibly produced.

Adults (Search-for-alternatives, and wm capacity for multiple models) can in principle construct alternative integrated models of the premises because they master Search-for-alternatives. Further, as they have a wm capacity for multiple models, they can base their answer on several, possibly all, of the integrated models obtained. However, adults are not immune to errors: they may sometimes make each error predicted for children and adolescents.

5.3. Mastery of Consistency-&-Equivalence

Mastery of Consistency-&-Equivalence occurs again in two steps. People with only a partial mastery comprehend the necessity of checking whether the two putative conclusions are consistent and equivalent, but they confound the two different tests. These subjects think that no valid conclusion (Nvc) follows both from two inconsistent models (correct deduction), and from two models that are consistent, but nonequivalent (incorrect deduction). Johnson-Laird and Bara (1984) noticed a number of these subjects, who tended to give a 'Nvc' answer any time they were presented with a multiple model syllogism. People with full mastery of Consistency-&-Equivalence not only understand its necessity, but are also able to distinguish between the meaning of the two tests. Because models that are consistent but non equivalent admit a valid conclusion compatible with them, fully competent subjects will limit the 'Nvc' answer only to the cases of inconsistency. Here too, the wm capacity is a correlated cognitive resource: in fact, subjects have to keep two models in working memory, in order to evaluate them.

5.4. Constraints in the program

The constraints imposed by mastery of Consistency-&-Equivalence and cognitive resources together, support the following predictions about the performance of subjects belonging to different age groups (Fig. 9):

Children (no mastery of Consistency-&-Equivalence, and wm capacity for one model) will tend to give a conclusion based on the first integrated model of the premises.

Adolescents (partial mastery of Consistency-&-Equivalence, and wm capacity for two models) will tend to give 'Nvc' responses to multiple model problems. As their mastery improves, the percentage of erroneous 'Nvc' answers decreases with the age.

Adults (full mastery of Consistency-&-Equivalence, and wm capacity for multiple models) are in principle able to give correct conclusions to both valid and invalid problems. However, as some adults confound consistency with equivalence, they may give 'Nvc' answers to valid problems.

In sum, our computational model assumes that some constraints may affect the reasoning process at the Falsification level. The program, based on these assumptions, enables us to predict the correct and erroneous responses of subjects belonging to different age groups. The model gets all and only the correct answers when no constraint is imposed on it.

The following three sections are devoted respectively to an analysis of syllogistic, relational and propositional reasoning. We first evaluate the predictions worked out by UNICORE, and then compare them with the experimental data on the human performance. Syllogistic data come from Bara et al. (1995), as they are the only developmental data in the literature which cover the whole set of 64 syllogisms. Relational data are the classical ones by De Soto, London and Handel (1965). In the propositional case, data in the literature do not allow to compare connectives in different tasks; thus, we devised the new experiment described in Section 8.

6. Syllogistic reasoning

In this section we validate UNICORE through two dimensions: first, we show that it explains the difference between competence and performance in syllogistic reasoning, accounting for both correct and erroneous deductions. Second, the model reproduces the developmental trend revealed by subjects of different ages when reasoning with syllogisms.

MMT gives an account of the difference in difficulty among syllogisms on the basis of both the number of integrated models and the specific mental operations required. For each of the possible forms of the syllogistic premises there exist many possible representations. We have chosen the following ones:

A. universal affirmative: 'all of the architects are bankers'

```
a — b
    b
```

I. particular affirmative: 'some of the architects are bankers'

```
a — b
a - - - b
```

E. universal negative: 'none of the architects is a banker'

```
 a
―――――――
      b
```

O. particular negative: 'some of the architects are not bankers'

```
 a
―――――――
 a - - - b
```

There exist four possible arrangements of the terms inside the premises, resulting in four possible figures (see note 1). The figure of the premises may require a Rearrange operation in order to make adjacent the middle terms *b* and to make Integration possible. In particular, the 1st figure AB-BC requires no rearrangement and UNICORE predicts A-C as the form of conclusions. The 2nd figure BA-CB requires a Reorder (result: CB-BA), and the model predicts C-A as the form of conclusions. The 3rd figure AB-CB requires an Inversion of the second premise representation (result: AB-BC), and UNICORE predicts A-C conclusions. The 4th figure BA-BC requires an Inversion of the first premise representation (result: AB-BC) and the model predicts A-C conclusions. In this last figure, since the reasoner might also reorder the premises (intermediate result: BC-BA), and then invert the first premise (result: CB-BA), C-A conclusions are possible, although not preferred.

The number of representations resulting from the integration and the figure of the premises allows the prediction of different levels of difficulty for syllogisms. Indeed, Bara et al. (1995) demonstrated that the percentages of correct conclusions drawn from syllogistic premises is a function of both the number of integrated models of the premises and the figure of the premises. Such results hold for subjects ranging all over the age groups. For an exhaustive comparison through the all syllogisms, the Appendix shows both UNICORE's predictions and experimental data for each of the individual syllogisms (28 syllogisms for the children, and 64 for the adolescents and the adults).

Let us consider some syllogistic inferences in order to analyze how UNICORE works in predicting the difficulty of the reasoning processes. We focus on both the easiest and the hardest of the 28 syllogisms presented by Bara et al. (1995) to children, adolescents and adults.

First, consider the easiest syllogism (82% of correct responses over all subjects):

[1] All of the athletes are baritones
None of the cadets are baritones

Therefore, none of the athletes are cadets

It requires an Inversion of the model of the second premise in order to permit the integration with the model of the first premise. The integration produces an integrated model of the premises. Falsification does not succeed in finding an alternative integrated model of the two premises; therefore, the correct conclusion 'None of the athletes are cadets' or vice versa, can be based on a single representation. In Fig. 10 we illustrate the mental processes involved in reasoning with the premises in [1]. For this figure only, we have detailed each step, pursuing the completeness of the description. Both the successful and the failed attempts to manipulate the models (continuous and dashed lines, respectively) are specified and numbered.

The figure reads as follows:

a) Integration phase

The reasoner attempts to integrate the models of the premises, but Match fails {1}, and it calls Rearrange. Rearrange calls a Reorder {2}, but Match fails again {3}. Then, Rearrange calls an Inversion of the second premise {4}, and produces a model that can be integrated with the model of the first premise: Match succeeds {5}.

b) Conclusion phase

Now Extractor produces the first result model {6}. At this point, Falsification is invoked.

c) Falsification phase

A new Match {7}, a Flesh-out {8} and a Token-moving {9} procedure are attempted in the mentioned order. None of them succeeds in producing an integrated model of the premises which is different from the one previously obtained.

d) Response phase

The current result model is eventually passed to the Response phase, and the Generator-of-responses gives the verbal output {10}.

Consider now the hardest of the syllogisms (5% of correct responses over all subjects):

[2] None of the bandits are aviators
Some of the bandits are contortionists

Therefore, some of the contortionists are not aviators

In syllogism [2], the result of Integration is a first integrated model where none of the aviators are contortionists (Fig. 11).

Falsification procedure applies Token-moving. The result is a second integrated model where some of the aviators could be contortionists; the model supports the conclusion 'Some of the aviators are not contortionists' or vice versa. Both conclusions are also consistent with the first integrated model. Finally, the third model - where all the aviators could be contortionists - rule out the previous conclusion 'Some of the aviators are not contortionists'. In order to draw the correct conclusion 'Some of the contortionists are not aviators' the reasoner must take into account the three integrated models of the premises, two of which are produced by Token-moving.

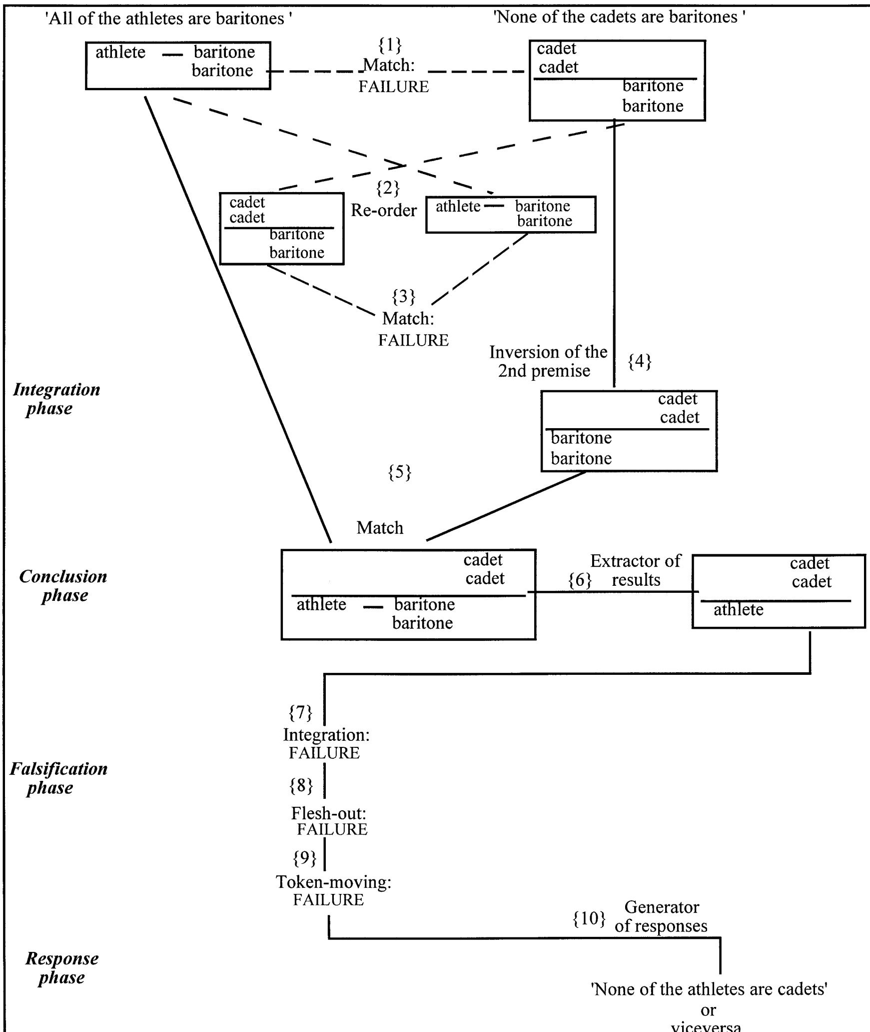


Fig. 10. Reasoning with the syllogistic premises 'All of the A are B. None of the C are B' [1]. The dashed lines represent the failed attempts to manipulate models; the continuous lines the successful ones. All of the attempts are numbered.

To sum up, the two syllogisms require a different number of representations and manipulations. Syllogism [1] requires an inversion of the 2nd premise and the evaluation of one integrated model of the premises; syllogism [2] requires an inversion of the 1st premise and the evaluation of three integrated models of the premises. MMT accordingly predicts different performances in human subjects. They should perform well with syllogism [1]. In particular, children are expected to perform as well as adolescents and adults. On the contrary, the correct conclusion to [2], 'Some of the contortionists are not aviators', should be quite hard to draw even for adult subjects. But let us detail the predictions for [2] according to each age group.

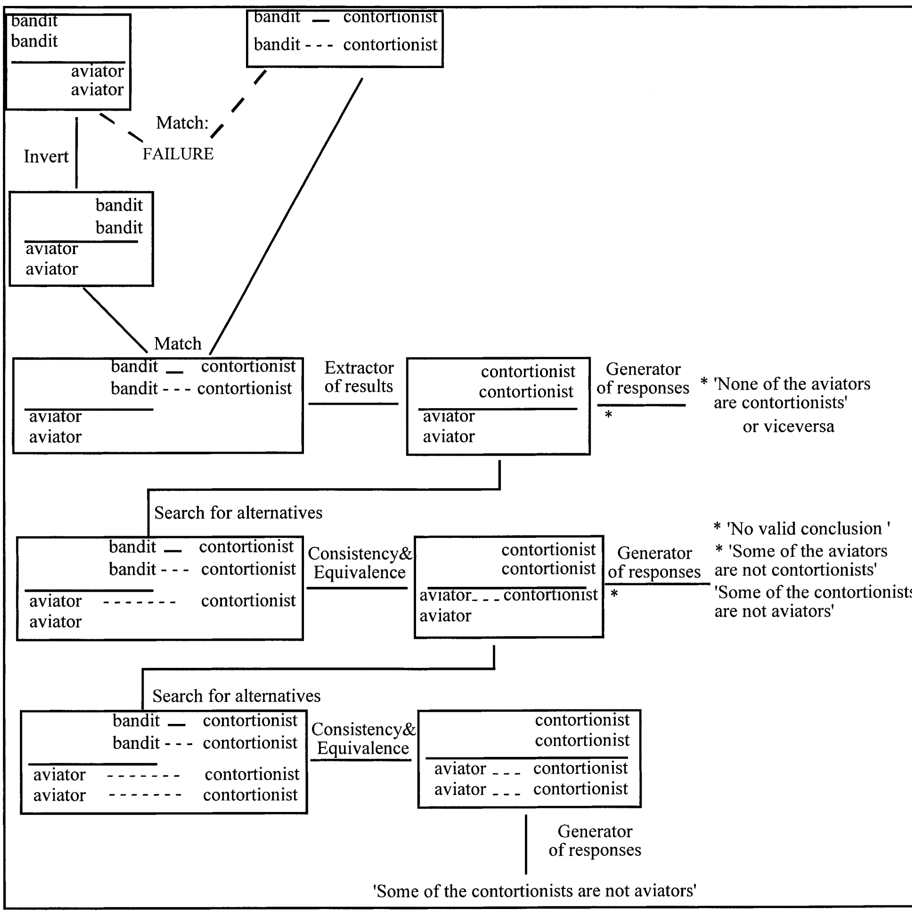


Fig. 11. Reasoning with the syllogistic premises 'None of the B are A. Some of the B are C' [2]. * marks both premature exits from the reasoning process and errors.

Children

We expect that, due to their lack of mastery of Falsification and to their small working memory capacity, children will base their response just on the first integrated model of the premises, answering 'None of the aviators are contortionists' (the A-C conclusion is preferred), or vice versa.

Adolescents

Adolescents should be able to produce a second integrated model alternative to the one previously obtained. Thus, they may draw either the erroneous conclusion 'Some of the aviators are not contortionists' or the correct one 'Some of the contortionists are not aviators'. In fact, they partially master Falsification, and the individual differences are due to the capacity of the working memory of each subject.

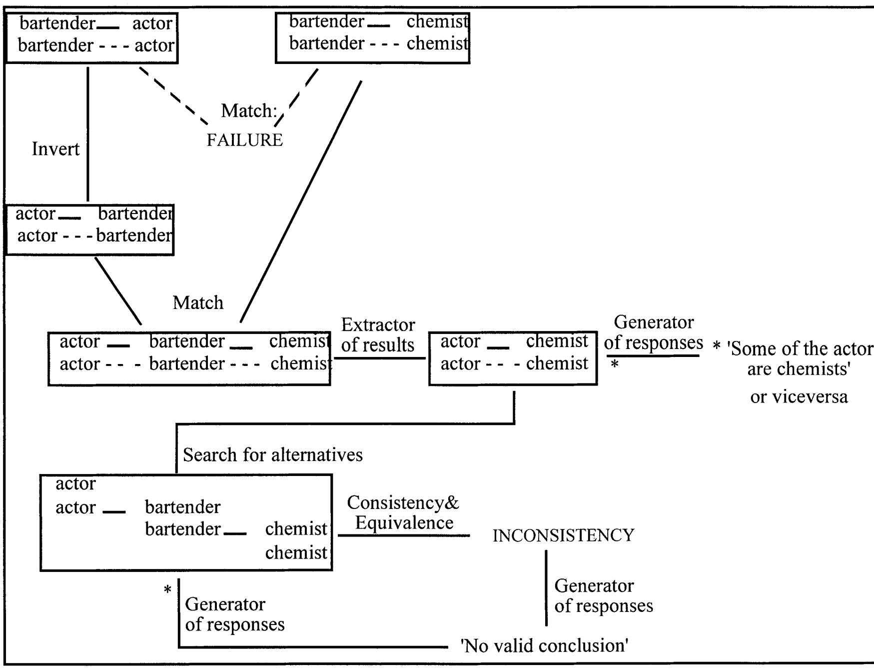


Fig. 12. Reasoning with the syllogistic premises 'Some of the B are A. Some of the B are C' [3]. * marks both premature exits from the reasoning process and errors.

Adults

Because of their working memory capacity for multiple models and full mastery of Falsification, adults are the most likely to produce the correct answer: 'Some of the contortionists are not aviators'. For both adolescents and adults we also predict the erroneous response 'No valid conclusion', which is generated by the subjects who treat as inconsistent two models which are qualitatively different.

Let us now consider the reasoning process involved in reasoning with a multiple-model problem without a valid conclusion. Here we analyze an invalid syllogism of medium difficulty (37% of correct conclusions over all subjects):

[3] Some of the bartenders are actors
Some of the bartenders are chemists

Therefore, no valid conclusion

After the Inversion-of-the-first-premise, Integration succeeds (see Fig. 12).

The result is a model where some of the actors are chemists (the A-C conclusion is preferred) or vice versa. We expect that children will give this type of erroneous response. Then, Falsification is applied; a new Integration followed by a Token-moving produces a representation where nothing can be asserted about the relation between actors and chemists. In order to draw the correct conclusion 'No valid conclusion', the reasoner has to build the

Table 1
Percentages of correct conclusions with the syllogisms [1], [2] and [3]

		Age groups (n = 20)			
	Syllogisms	9–10	14–15	>21	Total (N = 60)
[1]	All A are B No C are B	80	90	75	82
[2]	No B are A Some B are C	0	5	10	5
[3]	Some B are A Some B are C	15	40	55	37

second integrated representation: we expect that adults are more likely to produce all the integrated models and to base the conclusion on them. It is interesting to note that the correct response ‘No valid conclusion’ is also predicted for those subjects who do not master Consistency-&-Equivalence. Often, in invalid syllogisms the correct ‘Nvc’ conclusion may be reached *via* shortcuts.

Now, we compare our predictions about the three syllogisms with the detailed data presented in Bara et al. (1995). Global percentages of correct conclusions to the three syllogisms are in Table 1.

UNICORE’s predictions, together with the experimental data, are illustrated in Tables 2, 3 and 4, which respectively refer to syllogisms [1], [2] and [3]. Tables 2, 3 and 4 are organized as follows. First, consider the leftmost column. Each entry corresponds to a type of response given to the syllogism. The correct responses are printed in capital letters. Responses above the correct one are the erroneous responses which are predicted by UNICORE. They are listed in the predicted order; they are based on the model representations that the Extractor may read out of each integrated model of the premises, during the reasoning process. Responses below the correct response are unpredicted responses that were given by at least two subjects. Thus, percentages in the tables do not necessarily sum up to 100%.

The second, third and fourth columns respectively correspond to the percentages of types of response given by human subjects in the 9–10, 14–15 and over 21 age groups (20 subjects in each group). Columns five, six and seven correspond to the model’s predictions about the possibility that a type of conclusion is drawn (+) by artificial subjects. The degree of expectation of a response in a specific age group is reflected by the respective number of crosses. As the number of crosses ranges from 1 to 3, they represent a crude simplification of the data. In fact, UNICORE’s predictions are based on an ordinal scale. Predictions are meant to hold both within and between age groups. Thus, they are satisfactory when they get close to the data. A different sign (x) is used for predictions concerning the adult group, as a reminder that predictions for them are only qualitative. By this different notation, we intend to emphasize that adult subjects fully express the competence level, but they are not immune from logical errors.

The syllogisms we utilize in the following contain the same information as explained in the Appendix. Predictions for syllogism [1] are confirmed (see Table 2). The only surprise is that C-A is the favorite form of conclusion for adolescents and adults.

Table 2
Types and percentages of responses given by experimental subjects and artificial subjects to the syllogistic premises in [1]

All A are B No C are B	Age groups (n = 20)			Artificial subjects		
	9–10	14–15	>21	children	adolescents	adults
NO A ARE C	65	35	25	++	++	x
NO C ARE A	15	55	50	+	+	x
Nvc	10		25			

Predictions for syllogism [2] are also confirmed (see Table 3). Almost nobody infers the correct conclusion from the premises, and the most common erroneous response is based on the first integrated representation of the premises.

Predictions are also confirmed for syllogism [3], where the most common erroneous responses are based on the first integrated representation of the premises (see Table 4).

Let us now offer a more comprehensive analysis of one model and multiple model syllogisms. The analysis refers to the data in the Appendix. For a comparison involving the three groups of subjects we shall consider the 28 syllogisms also presented to children. Table 5 shows the percentages of correct responses given by children, adolescents and adults to one model syllogisms. Children's performance does not significantly differ in comparison with that of the adolescents and the adults. Such data confirm the general prediction of MMT according to which one model problems are easy to solve, even for young subjects.

Multiple model problems are more difficult to solve. Table 6 shows the percentages of responses - either based on the first integrated model or on further integrated models - given by the same groups of subjects to multiple model syllogisms. Erroneous responses are very common; many are based on the first possible model of the premises. However, with the increase of age, there is also an increase in the number of responses based on further integrated models. Again, such results confirm a prediction of MMT. We conclude that UNICORE accounts for the performance of subjects belonging to different age groups when reasoning with syllogisms.

UNICORE's predictions on the 64 classical syllogisms are consistent with previous predictions by Johnson-Laird and Bara (1984). However, four syllogisms are an exception.

Table 3
Types and percentages of responses given by experimental subjects and artificial subjects to the syllogistic premises in [2]

No B are A Some B are C	Age groups (n = 20)			Artificial subjects		
	9–10	14–15	>21	children	adolescents	adults
No A are C	60	65	30	+++	++	x
No C are A	15		5	+	+	x
Nvc	5	5	50		+	x
SOME C NOT A		5	10		+	x
Some A are C	5	10				
Some C are A	15	5	5			

Table 4
Types and percentages of responses given by experimental subjects and artificial subjects to the syllogistic premises in [3]

Some B are A Some B are C	Age groups (n = 20)			Artificial subjects		
	9–10	14–15	>21	children	adolescents	adults
Some A are C	60	55	30	+++	++	x
Some C are A	15		15	+	+	x
NVC	15	40	55		+	x

In particular, syllogisms Iab-Abc, Aba-Icb, Aba-Ibc and Iba-Abc are multiple model problems in UNICORE, whereas they were previously considered as one model problems. This fact can be explained in terms of model construction from quantified premises. Indeed, the four syllogisms share a common feature: one of the premises is in the A (all) mood, the other is in the I (some) mood. Given the initial representations of A and I premises in the program, their integration always give rise to two possible models. The point is which sort of models the reasoner starts with, because it affects the reasoning process. We shall come back to this crucial issue in section 9.

7. Relational reasoning

In this section we shall illustrate how the model accounts for both correct and erroneous relational deductions. However, as relational problems have never been investigated in children and adolescents, we test our predictions on adults only.

Let us focus on a kind of relational problem usually known as *three-term series problems.* The premises of such problems describe objects in relation by using relational terms like ‘better than’ and ‘worse than’. People usually produce correct conclusions to these problems (see Huttenlocher, 1968), unless they are forced to act within a strict time limit. De Soto et al. (1965), for instance, carried out an experiment where subjects were given 10 seconds for each problem, and found that fewer correct answers were produced. De Soto and colleagues argue that people - when drawing relational inferences - rely on spatial representations or ‘thought models’ which they construct in some cognitive space (for an exhaustive review of the theories on relational reasoning see Evans, Newstead & Byrne, 1993). MMT takes up their claim.

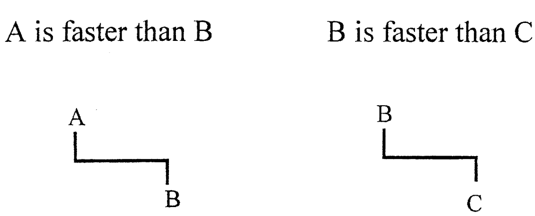
Table 5
Percentages of correct responses given by experimental subjects and artificial subjects to one model syllogisms. Such responses are based on the only possible model of the premises

	Age groups (n = 20)			Artificial subjects		
	9–10	14–15	>21	children	adolescents	adults
1st model’s conclusions	56	64	67	++	++	x
Unpredicted	44	36	33			

Table 6
Percentages of predicted responses (both erroneous and correct) given by experimental subjects and artificial subjects to multiple model syllogisms

	Age groups (n = 20)			Artificial subjects		
	children	adolescents	adults	children	adolescents	adults
Conclusions based on the 1st model	40	35	24	++	+	x
Conclusions based on further models	38	50	68	+	++	x
Unpredicted	22	15	8			

Thus, given a pair of relational premises such ‘A is faster than B; B is faster than C’, UNICORE assumes that the reasoners construct the following analogical representations:



This notation expresses the analogical structure of the representations involved in reasoning with comparative terms. It constitutes a novelty in comparison with Johnson-Laird (1983), whose proposal has the disadvantage of not allowing the left-to-right arrangement of the input tokens in working memory. Such an arrangement is meant to reflect an operative feature of working memory (see section 3). Besides, in comparison with the treatment proposed by Hunter (1957), our notation has the advantage of being content independent.

Predictions are made according to the difficulty of the rearranging operations that are invoked to integrate the models of the premises.

First, consider the relational premises:

[4]	A is faster than B	[5]	B is slower than A
	B is faster than C		C is slower than B

[4], which requires no rearranging of the premises (Fig. 13), is expected to be easier than [5], which requires the application of Reorder to the models of the premises (Fig. 14).

Second, consider the relational premises

[6] B is faster than A
B is slower than C

[6] is expected to be far more difficult than [5]. Indeed, inference [6] requires the inversion of the model of the first premise in order to produce an integrated model (Fig. 15).

Percentages of correct responses given by adult subjects to the three problems are

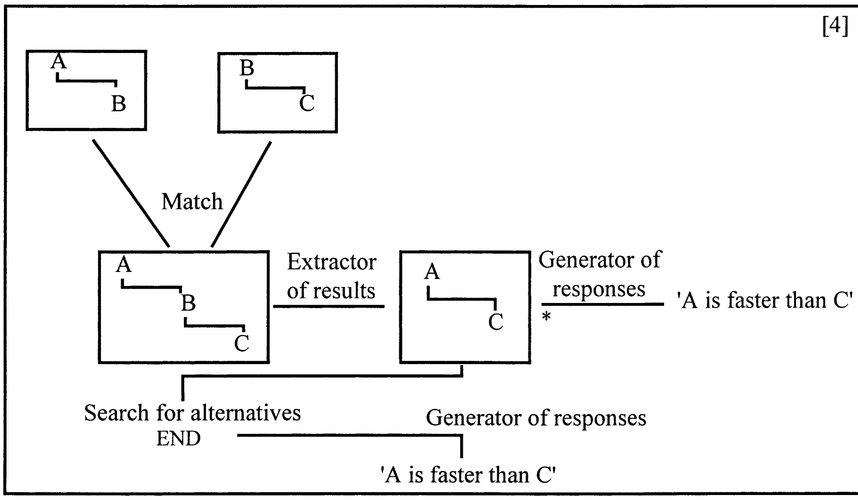


Fig. 13. Reasoning with the relational premises ‘A is faster than B. B is faster than C’ [4]. * marks premature exits from the reasoning process.

summarized in Table 7 (De Soto et al., 1965). Predictions are confirmed: inferring a conclusion from the premises in [4] is easier than drawing a conclusion from the premises in [5]; in turn, the latter problem is easier than problem [6]. We conclude that

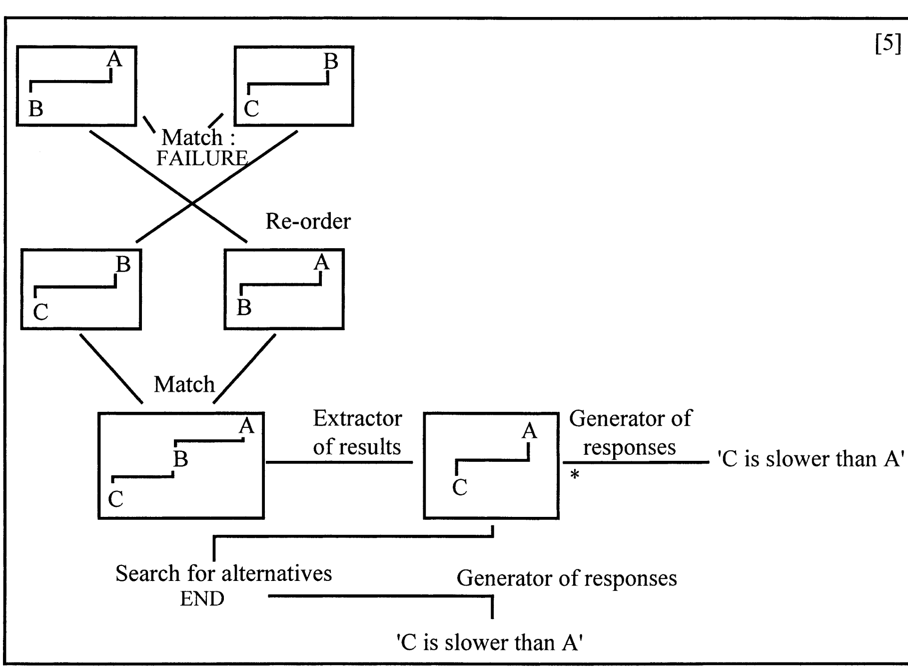


Fig. 14. Reasoning with the relational premises ‘B is slower than A. C is slower than B [5]. * marks premature exits from the reasoning process.

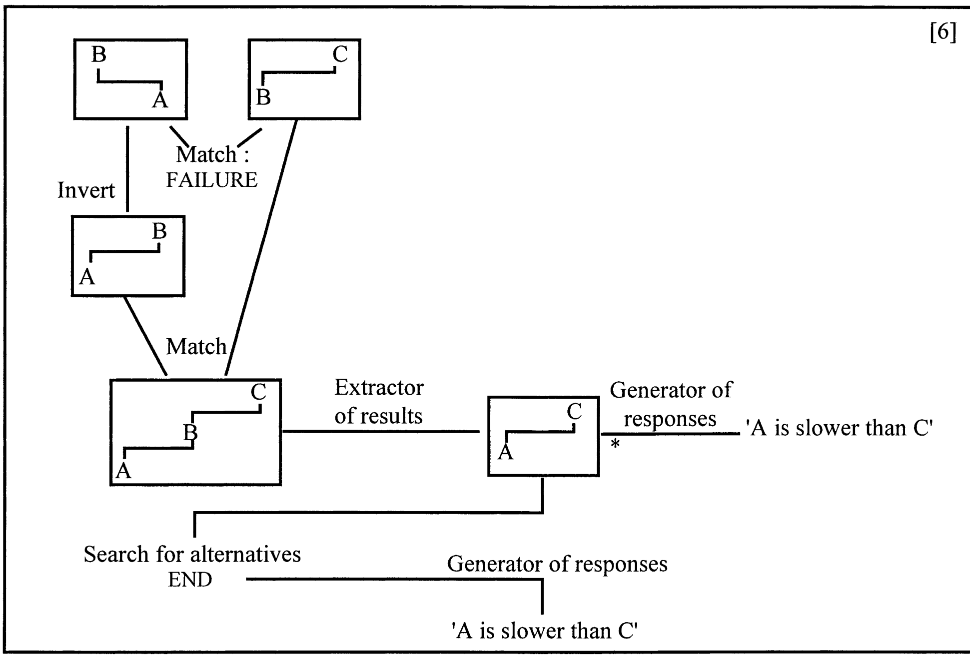


Fig. 15. Reasoning with the relational premises 'B is faster than A. B is slower than C [6]. * marks premature exits from the reasoning process.

UNICORE predicts human performance with the three-term series problems we have analyzed.

8. Propositional reasoning

In this section, UNICORE is validated along two dimensions. First, it accounts for the difference between competence and performance; in fact, the model accounts for both correct and erroneous propositional deductions. Second, the model explains the developmental trend

Table 7
Percentages of correct responses given by adult subjects and artificial subjects to the relational premises [4], [5], [6] with a 10 sec. time limit.

	Problems	% of correct responses (n = 117)	Artificial subjects
[4]	A faster B B faster C	61	+++
[5]	B slower A C slower B	50	++
[6]	B faster A B slower C	38	+

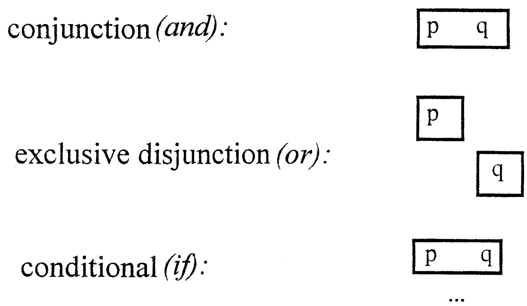
revealed by subjects of different ages when reasoning with connectives. A key point to note is that the comprehension of a connective occurs inside the context provided by a particular reasoning task. A specific novelty of UNICORE is to explain how the task in which the connective is involved affects the reasoning process, thus answering the criticism of Bonatti (1994), who argues that MMT presupposes but does not explain how context affects the reasoning process. No theory of propositional reasoning has previously accounted for such context-dependent meaning of connectives.

Experimental data show that connectives vary in difficulty: conjunction is handled by 2 year old children (Bloom, Lahey, Hood, Lifter & Fiess, 1980), disjunction is understood after 4 years (Johansson & Sjolin, 1975), and the conditional remains difficult even for 14 year old children, although around 5–6 years a clear improvement emerges (Amidon, 1976; Staudenmayer & Bourne, 1977). However, psychological researches have not thoroughly investigated the role of the type of task in determining the difficulty of a connective. An exception is Braine and Rumain (1981), whose major finding is that the disjunctive connective is more difficult in truth-value-judgment than in inference; this fact is considered as evidence for the existence of mental inference rule schemata. However, Braine and Rumain's explanation does not give a clear account of the mental processes involved in propositional reasoning; therefore it is impossible to make fine predictions on differences in difficulty among connectives in tasks. A further claim of Braine and Rumain is that their result runs against MMT, whose major assumption is that reasoners construct mental representations quite similar to truth tables. If MMT was right, they argue, truth-value judgments would be easier than inferences. But their conclusion is based on an incorrect comparison of the disjunctive connective 'or' in the two tasks they consider. Specifically, they test subjects' performance in truth-value judgment task by means of 4 trials, and subjects' performance in inference task by a single trial, chosen among the easiest ones.

To investigate more deeply the issues left open by their work, we carried out a new experiment. What we are skeptical about is the existence of an absolute gradient of difficulty among connectives. Our basic hypothesis, on the contrary, is that the context of application (task) affects the mastery of the connectives. Therefore, our aim is to account for differences in difficulty among connectives both within and among tasks. Previous work in the MMT paradigm has been concerned with the difference in difficulty among connectives in the inference task (Johnson-Laird, Byrne & Schaeken, 1992). We have investigated connectives in truth value judgment and action - along with the inference task. Thus, we provide an empirical extension of MMT to new domains, and to the never explored interrelation between connectives and tasks.

The experiment analyzes the comprehension of conjunction, exclusive disjunction and conditional in three different tasks: Inference (to derive a conclusion from two premises), Truth-value judgment (to judge whether an assertion is true or false with respect to a state of affairs), and Action (to act according to a request). We assume that even our youngest experimental subjects comprehend the essential meaning of the three connectives, conveyed respectively by the following models:

Nonetheless, we think that it is possible to envisage different levels of knowledge about the full meaning of the connective 'if' (see section 3). In particular, given a conditional assertion like 'If p then q', we claim that the meaning of the assertion as specified by the



co-occurrence of *not-p* and *not-q* is acquired later, whereas the meaning specified by the co-occurrence of *not-p* and *q* is almost unknown even to older subjects. The reason is probably the fact that this sort of backward inference is usually not experienced in daily life reasoning (Geminiani, Carassa & Bara, 1996).

The complexity of the experimental design reflects the necessity of investigating the three connectives as applied in quite different tasks. In order to make our predictions comprehensible, let us present the experiment.

8.1. The experiment on propositional reasoning

8.1.1. Experimental subjects

Twenty subjects (10 females and 10 males) from each of the following age groups: 3–4, 5–6, 8–9, 11–12 and over 21. They had no previous formal training in logic.

8.1.2. Design and Procedure

Subjects were told they were taking part in a test on the way people reason. They were presented individually with the test in a single session. The trials were presented in the following order: first, subjects dealt with the action and the inference trials (we adopted four randomizations balanced for number of subjects and sex), then with truth-value judgment trials (we adopted a different randomization for each subject). The reason is that in the pilot experiment we found that younger subjects got confused when dealing with truth-value judgment and inference mixed together: they sometimes tried to judge the premises of an inference, or to derive a conclusion from an assertion regarding a state of affairs in the case of truth-value judgment.

8.1.3. Materials

The materials consisted of one box and a series of toy animals (elephant, giraffe, lion, etc.), which we shall identify with the symbols *p, q,* and so forth. The general argument is the presence/absence of the animals inside the box. The trials consisted of instances of the three tasks.

1. Inference. The reasoner is presented with a first premise of the following type:

'If p is in the box, then q is in the box' (*if*)
'Either p is in the box or q is in the box' (*or*)

Then follows a second premise of the form:
a) ‘p is in the box’. b) ‘p is not in the box’. c) ‘q is in the box’. d) ‘q is not in the box’.
The subject is invited to draw a conclusion.

2. *Truth-value judgment.* One of the following situations occurs: either both entities p and q are in the box (TT - it’s true that p and q are in the box), or p is in the box and q is outside the box (TF), or p is outside the box and q is in the box (FT). The arrangement is plainly visible to the reasoner. Subjects have to consider each of the following judgments:
‘There is a p and a q in the box. True or false?’ (*and*)
‘If there is a p in the box then there is a q in the box. True or false?’ (*if-then*)
‘Either there is a p in the box or there is a q in the box. True or false?’ (*or*)

3. *Action.* The entity p is in the box and the entity q is outside the box. Both are plainly visible to the reasoner. One of the following requests is uttered
‘If a p is in the box, then put a q in the box’ (*if-then*)
‘Either there is a p is in the box or put a q in the box’ (*or*)

8.2. Predictions for propositional reasoning

8.2.1. Inference

Let us consider two classical inferences with conditionals that have different levels of difficulty, that is, *Modus Ponens* and *Affirming the Consequent.* The two inferences differ in the number of model representations and the type of manipulations required. On the one hand, in the case of *Modus Ponens* the correct response is supported by the only possible integrated representation of the two premises (Fig. 16).

On the other hand, *Affirming the Consequent* requires taking into account the two possible integrated models of the premises: the first one results from the first Match, and the second one results from a new Match between the fully explicit representation of the first premise and the model of the second premise (Fig. 17).

We expect *Modus Ponens* to be drawn by almost the totality of subjects; the model does not predict erroneous outcomes. *Modus Ponens* should be easier than *Affirming the Consequent,* for which UNICORE predicts the correct ‘No valid conclusion’ response, but also the erroneous response ‘p’, which is based on the first integrated model.

After the analysis of the conditional connective in the Inference task, we analyze conjunctive and disjunctive connectives in Judgment and Action tasks. As conjunction in Action is meaningless from a pragmatic point of view, and it has therefore been excluded from the experiment, we focus on predictions concerning conjunction in the Judgment task, and disjunction in the Action task.

8.2.2. Conjunctive connective in truth-value judgment

Consider the following case: [7] The lion (p) is in the box and the giraffe (q) is outside. The following truth-value judgment is requested: ‘There is a lion and a giraffe in the box. True or false?’ The model of the premise consists of a token (r), which is in turn a model containing two elements (p, q). The model of the perceived state of affairs consists of a token (p), which is an element. The function Integration tries to match the tokens p and r, and it returns a failure (Fig. 18).

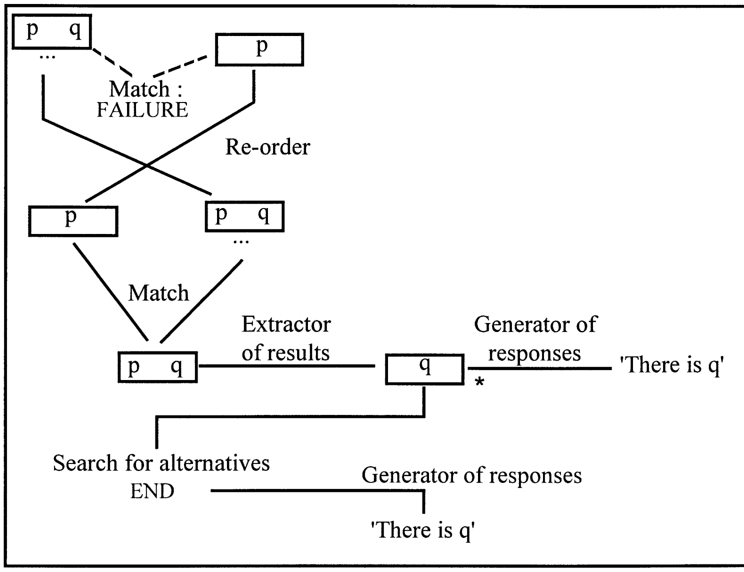


Fig. 16. *Modus Ponens*. Reasoning with the propositional premises 'If the dog is in the box (p), then the pig is in the box (q). The dog is in the box (p)'. * marks premature exits from the reasoning process.

We expect that almost all subjects will perform correctly with the conjunctive connective in such an occurrence.

8.2.3. Disjunctive connective in action

Consider the following case: [8] The tiger (p) is in the box and the elephant (q) is outside the box. The following request is uttered: 'Either there is a tiger in the box or put an elephant in the box'. For this task we need a particular type of token to represent an action to be performed (Fig. 19). The notation act (p) satisfies this requirement. The action functors (act) are a rough notation for complex objects of a theory of action, obviously outside the scope of the paper.

We expect [8] to be more difficult than [7] since for a successful integration, it requires applying Flesh-out token. Moreover, UNICORE predicts some difficulties for young subjects when dealing with the two models for disjunction. In particular, we predict that they have difficulties in fleshing out the negative token not-act(q) that is implicit in the first model for disjunction. As a consequence, we expect them to base their response on the act(q) token, and therefore to act.

8.3. Results of the experiment on propositional reasoning

The percentages of correct conclusions given to the *Modus Ponens* and *Affirming the Consequent* inferences are given in Table 8. Our prediction is confirmed: *Modus Ponens* is easier than *Affirming the Consequent* (25 of the 60 subjects perform equally with the two inferences, but the prediction holds for the remaining 35 subjects: $p = .5^{35}$).

The percentages of correct conclusions given to conjunction in Judgment [7] and disjunc-

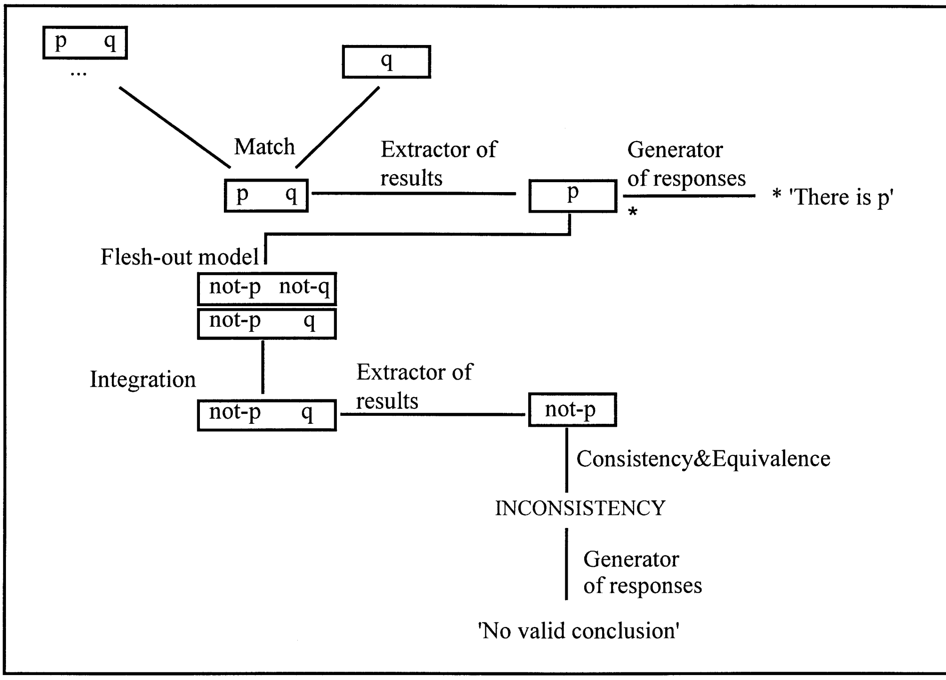


Fig. 17. *Affirming the Consequent*. Reasoning with the propositional premises 'If the bear is in the box (p), then the cow is in the box (q). The cow is in the box (q)'. * marks both premature exits from the reasoning process and errors.

tion in Action [8] are given in Table 9. Again, our prediction is confirmed: [7] is easier than [8] (Wilcoxon Test: z = -3.68, $p < .0001$).

Comparisons between predictions and experimental data are summarized in Tables 10 (*Modus Ponens*), 11 (*Affirming the Consequent*), 12 (*Conjunctive connective in Judgment*) and 13 (*Disjunctive connective in Action*). The Tables are organized according to the criteria

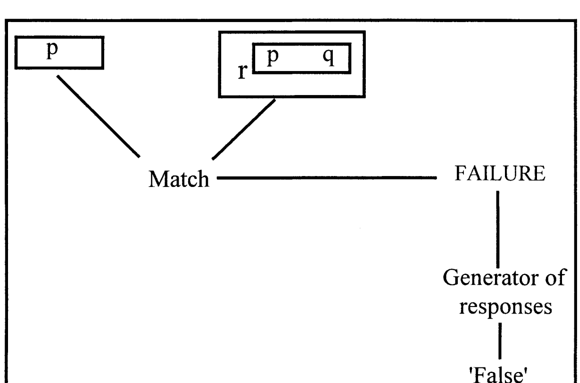


Fig. 18. Judging as true or false the assertion 'There is a lion (p) and a giraffe (q) in the box', with respect to a state of affairs where only a lion (p) is in the box [7].

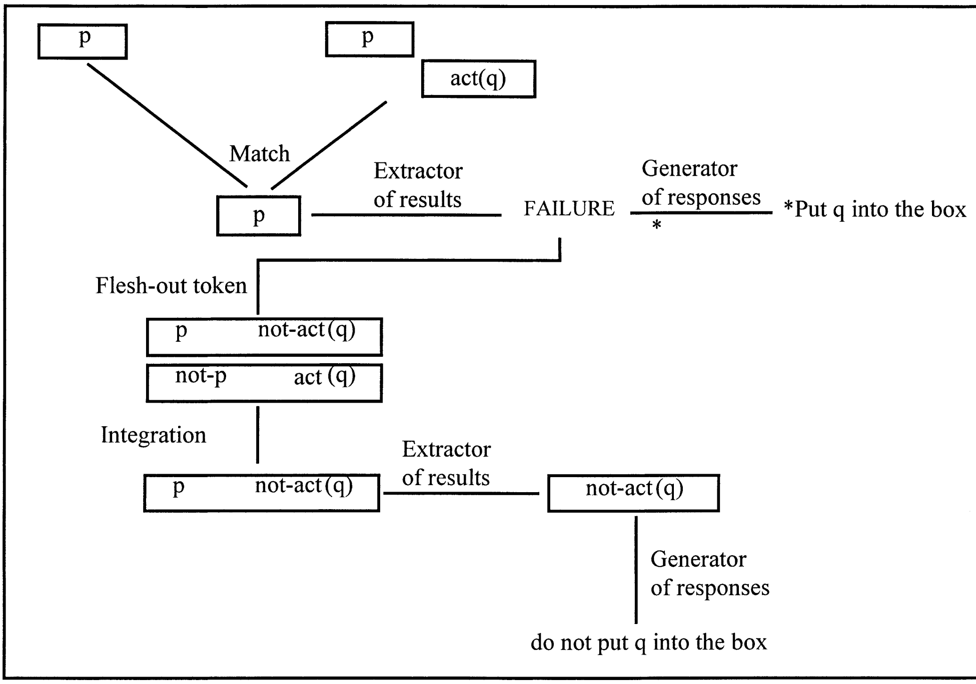


Fig. 19. Acting the request 'Either a tiger (p) is in the box or put an elephant (q) in the box', with respect to a perceived situation where a tiger (p) is in the box [8]. * marks both premature exits from the reasoning process and errors.

introduced in the syllogistic session. Here is a brief summary. Each entry in the leftmost column corresponds to a type of response given to the propositional problem. The correct responses are printed in capital letters. Responses above the correct ones are the erroneous responses predicted by UNICORE. They are listed in the predicted order. Responses below the correct responses are unpredicted responses that are given by at least two subjects. The second, third and fourth columns respectively correspond to the percentages of the types of response given by human subjects. Subject's ages were collapsed into three main categories: children (3–4 and 5–6 years old), adolescents (8–9 and 11–12 years old), adults (over 21 years old). Columns five, six and seven correspond to the model's predictions about the possibility that a type of conclusion is drawn (+) by artificial subjects. As for the model's

Table 8
Percentages of correct responses to *Modus Ponens* and to *Affirming the Consequent* inferences

Inferences	Age groups (n = 20)					Total (N = 100)
	3–4	5–6	8–9	11–12	>21	
Modus Ponens	5	25	15	55	100	40
Affirming the Consequent	0	0	0	0	25	5

Table 9
Percentages of correct responses to problems [7] and [8]

Occurence of the connective		Age groups (n = 20)					Total (N = 100)
		3–4	5–6	8–9	11–12	>21	
[7]	Situation: p Assertion: 'p and q'	40	55	95	100	100	78
[8]	Situation: p Assertion: 'p or put q'	20	20	50	75	95	52

predictions in syllogistic inferences, the degree of expectation about the appearance of a response in a specific age group is reflected by the number of crosses. A different sign is used for the adult group (x), where subjects are supposed to have full access to the competence level, even if they are not immune from logical errors.

Predictions are only partially confirmed for *Modus Ponens*: many children and adolescents produce erroneous responses not predicted by our model (see Table 10).

Predictions for *Affirming the Consequent* are partially confirmed (see Table 11). The most common erroneous response is based on the first integrated model. However, the majority of adolescents prefer the conclusion supported by the second integrated model (not-p) alone, without checking its consistency with the previous model. Currently, UNICORE - as all the computer programs devised inside the mental model paradigm - does not allow for such an effective possibility. Indeed, Bucciarelli and Johnson-Laird (1999) have found evidence of the fact that, although sometimes reasoners construct all the possible models of the premises, they may err by drawing a conclusion which is based on the latter model they have produced. Thus, this constraint will be the next one to be introduced in UNICORE.

The results are only in part consistent with those in Rips (1994), but this fact can be explained by the type of task employed by Rips, that is an evaluation task, while we use a production task. Specifically, Rips' subjects are required to judge a conclusion as true or false, while our subjects have to draw a conclusion on their own. A genuine surprise comes from the inversion of the conclusion (p/not-p) shown by the 11–12 years old subjects, who seem to go back to a previous stage in performance. Bara et al. (1995) have already noticed a similar decrease in adolescents' performance, but we have no explanation for this finding.

Let us now consider the predictions and experimental data for connectives in tasks. Table

Table 10
Types and percentages of responses given by experimental subjects and artificial subjects to *Modus Ponens* inference

Modus Ponens	Age groups (n = 20)					Artificial subjects		
	3–4	5–6	8–9	11–12	>21	children	adolescents	adults
Q	5	25	15	55	100	+	+	x
not-q		30	85	40				
p	55	35						
say nothing	15							

Table 11
Types and percentages of responses given by experimental subjects and artificial subjects to *Affirming the Consequent* inference

Affirming the Consequent	Age groups (n = 20)					Artificial subjects		
	3–4	5–6	8–9	11–12	>21	children	adolescents	adults
p	25	20	15	60	65	++	++	x
NVC					25		+	x
not-p	5	30	85	40	10			
q	30	25						
say nothing	15							
I don't know	25	15						

12 summarizes the predictions and data for *conjunctive connective in Judgment.* Almost all subjects beyond childhood perform correctly, as predicted by the model.

Table 13 concerns *disjunctive connective in Action.* As predicted, the difficulty of fleshing out the not-act(q) token leads young subjects to base their response on the act(q) token, and therefore to act. On the other hand, with increasing age the frequency of correct performance increases.

The results on connectives deserve some consideration, in particular those erroneous outcomes not predicted by the computational model. Note that the erroneous response 'p' given by children to *Modus Ponens* (Table 10) is in contradiction with the competence expressed by the Conclusion phase, because it is a plain repetition of one of the premises. The crucial point is that the response 'p' can be drawn from the integrated model representation as well as the response 'q', and the Extractor-of-results is responsible for the erroneous selection of information. The same can be said of the erroneous response 'q', given by the same subset of subjects to *Affirming the Consequent* inference (Table 11). Such a response may be the result of an erroneous selection of information, by the Extractor, from the first integrated representation of the premises.[9] As the Extractor behaves according to the task at hand, we argue that the erroneous performance of children might be caused by poor knowledge of the task.

In summary, since our model allows us to make detailed predictions about the performance of subjects dealing with connectives in tasks of different kinds, we claim that MMT has the power to explain the meaning of connectives in all the relevant contexts. What our results show is that the grasp of connectives does not depend upon absolute comprehension. On the contrary, subjects master each connective within the context where it is utilized (in our case, the task). For this reason, our developmental predictions are derived not from abstract differences, but from contextualized uses of the connectives.

Table 12
Types and percentages of truth-value judgements given by experimental subjects and artificial subjects to conjunctive connective in Judgment [7]

[7] Situation: p Assertion: 'p and q'	Age groups (n = 20)					Artificial subjects		
	3–4	5–6	8–9	11–12	>21	children	adolescents	adults
FALSE	40	55	95	100	100	+	+	x
True	60	45	5					

Table 13
Types and percentages of responses given by experimental subjects and artificial subjects to disjunctive connective in Action [8]

[8] Situation: p Assertion: ‘p or put q’	Age groups (n = 20)					Artificial subjects		
	3–4	5–6	8–9	11–12	>21	children	adolescents	adults
Act	75	75	50	25	5	++	+	x
NOT ACT	15	15	40	65	90	+	++	x
PUT Q TAKE AWAY P	5	5	10	10	5			

9. Conclusions

Our model is an attempt at a unified computational program of deductive reasoning in the spirit of the mental model paradigm. First, we have presented an ontology which constitutes a formal foundation for the Theory of Mental Models. Second, we have illustrated our computational model of deductive reasoning based on mental models. Finally, we have successfully compared the performance of the human subjects of different ages versus the artificial subjects of a corresponding level of development.

Let us now explore the limits of the model.

i. UNICORE does not take into account the constraints on model construction, which is treated as error-free. In particular, we assume that, given a premise, everybody constructs a correct mental representation, that is, a representation where all the input information is maintained in the analogical model. Obviously, it is quite unlikely that everybody, from the youngest age, possesses a complete knowledge of quantifiers, connectives and relational terms. Thus, it is highly probable that some subjects construct an incomplete or incorrect model representation of the input. If they cannot form a proper representation, they will be unable to draw the correct conclusion, even if their deductive capacity were impeccable.

Further, it is possible that people differ according to the sort of (correct) representations they individually build starting from the premises of a reasoning problem. If this is the case, the reasoning process might be affected depending on the different sorts of representations (see section 6). The present version of the model undervalues the relation between interpretation and deductive performance. Indeed, as Bucciarelli and Johnson-Laird (1999) point out, human reasoners follow a great variety of strategies and interpretations when reasoning deductively, and a deterministic framework hardly captures the highly flexible human system. However, a unified computational model of deduction aims to shape the mental processes involved in any kind of deductive reasoning. The consequence is a wide perspective on deductive competence without an attempt to account for individual variability.

ii. A second limit of our model is that it does not take into account the influence of the problem content of the reasoning process. The computational model does not take into account this factor, which has been carefully avoided in the experimental tasks we have employed. Further, because of its intended generality, the model does not take into account the constraints imposed by world knowledge and personal beliefs on the reasoning process. In a MMT perspective, these biases may enter the reasoning process any time the Integration or Falsification procedures result in a model representation. Oakhill, Johnson-Laird and Garnham

(1989) provide evidence in favor of the fact that, when reasoning with multiple model syllogisms, reasoners tend to accept the conclusion supported by the first integrated model if the conclusion is believable. A consequence is that the models of the premises are not always fully processed; complete processing is more likely to occur when the first integrated model supports an unbelievable conclusion, leading to an attempt to falsify the model itself.

iii. A third limit of the model is that it is concerned with syllogistic, relational and propositional reasoning, but not with reasoning with multiple quantifiers, spatial reasoning or temporal reasoning. As regards reasoning with multiple quantifiers, the computational model is consistent with the claims of Johnson-Laird, Byrne and Tabossi (1989). In a series of experiments they have found that adult subjects draw more correct deductions from the one model problems than from multiple model problems. In order to make our exposition as clear and succinct as possible we have decided not to attempt to reproduce their findings at this stage. However, the preliminary computational experiments we have carried out in this domain confirm that it ought to prove a straightforward extension of the model.

Our computational model does not reproduce the reasoning steps involved in spatial reasoning. Indeed, we assume that, in order to draw spatial inferences, the reasoner relies on spatial maps (Bower & Morrow, 1990). MMT claims that subjects build up model representations from spatial descriptions and, as spatial descriptions need to make references to the same set of entities in order to be integrated, the difficulty is predicted by the number of models (dictated by the number of referential sentences) and the continuous, semicontinuous and discontinuous referentiality. Given the matricial character of model representations, UNICORE is ontologically suitable to represent spatial arrays of objects. However, we believe spatial reasoning would require a not trivial extension of the program.

A recent development of MMT gives an interpretation of temporal reasoning along the same line of the spatial domain (Schaeken, Johnson-Laird & d'Ydewalle, 1996). Again, in UNICORE, the analogical structure of the token matrix and the character of the tokens can provide a useful representational medium for investigating the temporal domain.

Further, UNICORE does not deal with meta-logical problems, that is problems that explicitly concern matters of truth and falsity.[10] Johnson-Laird and Byrne (1990) claim that meta-logical thinking depends on two principal components:

a) The procedure for making ordinary deductions. In particular, in dealing with meta-logical problems the procedure carries out propositional deductions.

b) A high level component that uses various strategies in order to put the deductive mechanism to work. Indeed, the mechanism is used as a subcomponent. The high level component supports the ability to reflect on deductive problems.

Although in principle compatible with our current model, the implementation of the high level component would require a significant development.

Notwithstanding the explicit limits of the present research, the model described in this paper represents the core of a unified theory of deductive reasoning. People may - or may not - have mental processes dedicated to reasoning. Some authors reject the existence of such a specific mechanism. E.g., Ford (1994) and Polk and Newell (1995) argue that the manipulation of the verbal form alone may account for the ability of drawing correct syllogistic inferences. In direct contrast MMT implies that people have mental processes that procedurally express the falsification principle.

The present work is a demonstration in favor of the feasibility of a central mechanism devoted to deductive reasoning; such a mechanism may be computationally and experimentally investigated. Our model is a systematization of this idea, and it is aimed at providing a unified account of rationality which allows one to make predictions, rather than merely to describe experimental data *post hoc*. In particular, a unified theory of cognition enables one to see how multiple constraints jointly affect the reasoning process:

i. constraints that determine differences in difficulty among reasoning problems;

ii. constraints that may influence the reasoning process at different stages of development.

Our computational model is grounded on the assumption that it is crucial to explain how the ability to reason develops. Indeed, we claim that the competence of an adult system can be accounted for only if we are able to reconstruct how it develops with age. A developmental model has more explanatory power than a steady state model since it must incorporate further features, which are typical of the developing system. Our model is consistent with some constraints on human cognitive processes: these are working memory capacity and degree of mastery of Falsification, that influence the emergence of the logical competence with increasing age. The model allows detailed predictions about the performance of subjects belonging to different age groups in deductive reasoning tasks.

Throughout the paper, we have focused on the central role of the Millstone, because it represents the core of the unified mechanisms for reasoning. From an architectural point of view, however, deduction is the fruit of a continuous interaction between the Millstone, the Extractor-of-results, and the Consistency-&-Equivalence. As the latter two behave in different ways depending on the task at hand, they guarantee the link between the deductive processes and the context within which they occur. Moreover, the interface between the Millstone and the environment, represented by the Interpreter and the Generator-of-responses, is sensitive to the system's goals. In our model the type of task counts as a pragmatic factor which influence the model manipulation process. In taking into account the major components of deduction, the accuracy of the details has been sacrificed to the balance of the global framework. The map is necessarily approximate: at times, large scale attempts become useful, and almost compelling.

Notes

1. We follow a suggestion by Daniel Schwartz.
2. A syllogism consists of two premises and a conclusion. They occur in one of four *moods* shown here with their customary mnemonics (*AdfIrmo,* 'I affirm', and *nEgO,* 'I negate'):
   All A are B (A: universal affirmative)
   Some A are B (I: particular affirmative)
   No A are B (E: universal negative)
   Some A are not B (O: particular negative)
   To support a valid conclusion, the two premises must share a common term (the so-called *middle* term), and hence the premises can have four different arrangements

(or *figures*) of their terms:

1. A-B B-C 2. B-A C-B 3. A-B C-B 4. B-A B-C

3. Propositional reasoning is reasoning with propositions and connectives such as *and, if-then, only if, but, or* and *not.* A typical example is the inference called *Modus Ponens* which involves the if-then (conditional) connective. Given two premises like:
   If it is not raining, then I go for a walk.
   It is not raining.
   the meaning of the connective allows the derivation of the conclusion:
   I go for a walk.
4. A three-term series problem consists of two premises and a conclusion. Each premise refers to two terms linked by a transitive relation. One term is shared by the premises and allows one to draw a conclusion regarding the relation between the other terms. An example of a three-term series problem is:
   Albert is taller than Bernard.
   Bernard is taller than Christine.
   Therefore, Albert is taller than Christine.
5. A spatial deduction consists of two or more premises and a conclusion. The premises describe spatial relations between entities. The relation between entities whose relation is not stated in the premises can be inferred since each premise refers to at least one entity mentioned in another. An example of spatial deduction is:
   The glass is in front of the dish.
   The spoon is on the left of the dish.
   The fork is in front of the spoon.
   Therefore, the fork is on the left of the glass.
6. Stenning and Oberlander (1995) classify the mixture of linguistic and graphical representations as *limited abstraction representational systems.*
7. Evans (1993) and Platt and Griggs (1993) criticize MMT, noting that it does not explain how the process of fleshing out the implicit information is triggered. Here, we make it explicit that the process of fleshing out is triggered by a failure either in the Integration or in the Falsification phase. Moreover, both the knowledge of the meaning of the quantifiers, the relational terms and the connectives, and the mastering of the task support the appropriate rearrangement in expert subjects.
8. The careful reader will have certainly noticed that the task of inference in propositional reasoning is quite similar to the syllogistic and the relational ones. In fact, the three forms of reasoning involve the assertion of new information with respect to the premises. Actually, the Extractor behaves similarly in the three cases, by deleting the information that is common to the two input premises and by keeping the remaining information.
9. Constraints imposed on the reasoning process at the Extractor-or-results level would also account for some of the erroneous responses given by children to the syllogistic inferences analyzed above. Indeed, the 21% of the erroneous responses not accounted for by our model state the relation of the middle term *b* either with the term *a* or *c.*
10. A typical problem is the knight/knave problem. It concerns a realm where there are only two sorts of persons: knights, who always tell the truth; and knaves, who always lie. One of the inhabitants named George says, ‘Both Martha and I are knights’: but

when we consult Martha, she says ‘George is a knave.’ Is George a knight or a knave? What about Martha? The solution is that George is a knave and Martha is a knight.

Acknowledgments

The authors were supported in this research by M.U.R.S.T. of Italy (cofinanced project 2001, Deduction and induction).

We would like to thank the helpful colleagues who read and criticized earlier versions of this paper: Peter Ayton, Kyung Soo Do, Alan Garnham, Leonardo Lesmo, Gabriele Lolli, Jane Oakhill, Massimo Poesio, Patrick Sturt, Daniele Theseider Dupre‘. We also thank Nick Chater, James Greeno, Daniel Schwartz, Kurt VanLehn, and an anonymous referee, who provided us with comments and suggestions on the first draft of the manuscript.

We are particularly grateful to our longstanding friend and colleague, Philip Johnson-Laird, who has followed the story of this paper, with much invaluable advice. His open-mindedness does not mean that he entirely agrees with our interpretation of the mental model paradigm.

Appendix

Predictions of UNICORE for the 64 classical syllogisms, and corresponding responses given by human subjects (experimental data from Bara, Bucciarelli & Johnson-Laird, 1995).

The 64 pairs of syllogistic premises are shown, where each cell in Tables A1-A4 corresponds to a particular pair of premises and presents the frequencies of each sort of response that was given by at least 2 subjects. Thus, frequencies do not always sum up to 20. The correct responses are printed in CAPITAL LETTERS (E.g., NVC indicates a correct response. Nvc indicates an incorrect response). The erroneous responses predicted by the program are listed above the CORRECT one - in the order given by the computer model. Unpredicted answers are written below.

The data in the left-hand columns are the numbers of children (out of 20) giving each sort of response, and these data are only for the 28 syllogisms that the children received; the data in the middle columns are for the adolescents (out of 20); and the data in the right-hand columns are for the adults (out of 20). For those 36 syllogisms presented only to adolescents and adults, the left-hand columns are the data for adolescents and the right-hand columns are the data for adults. Each cell also shows whether a problem is a one-model syllogism (as shown by a single square in the cell), a multiple-model syllogism with a valid conclusion interrelating the end terms (two white squares in the cell), or a multiple-model syllogism with no valid conclusion interrelating the end terms (two black squares). In the tables, we maintain the classical MMT distinction in one and multiple models. However, in UNICORE syllogisms Iab-Abc, Aba-Icb, Aba-Ibc e Iba-Abc are multiple-model syllogisms, whereas in Johnson-Laird and Bara (1984) they were considered one-model syllogisms.

UNICORE’s predictions are not reducible to the distinction between one-model and multiple-model syllogisms, because it takes into account not only the number of models, but also the different manipulation procedures.

Table A1
Responses given by children, adolescents, and adults to syllogisms of figure AB-BC

	First premise					
A	A	□	ALL A ARE C	19	10	19
			Some A are C		5	
			No A are C		2	
A	I	□□	All A are C	2	9	
			SOME C ARE A	1		1
			SOME A ARE C	15	11	19
A	E	□□	No A are C		16	16
			No C are A		2	2
			Nvc			2
			SOME C ARE NOT A		-	-
A	O	■■	Some A are C		11	7
			Some C are A		1	1
			No A are C		1	1
			NVC			3
			Some A are not C		7	7
I	A	■■	All A are C	12	2	1
			Some C are A	2	4	1
			Some A are C	4	14	16
			NVC			2
I	I	■■	Some A are C	19	17	17
			NVC		2	3
I	E	□□	No A are C		14	15
			No C are A			3
			Nvc			1
			SOME C NOT A		1	1
			Some A are C		3	
I	O	■■	Some A are C		11	8
			Some C are A			1
			Some A are not C		7	5
			No A are C		1	
			NVC			6
E	A	□	NO A ARE C	1	16	17
			NO C ARE A	7	3	1
			All A are C	9		
E	I	□□	No A are C	8	8	6
			No C are A	1	2	2
			Nvc			1
			SOME A ARE NOT C		6	5
			Some A are C	6	3	5
E	E	■■	No A are C		19	12
			NVC		1	8
E	O	■■	No A are C		3	5
			No C are A			1
			Some A are not C		3	1
			NVC			8
			Some A are C		12	4
O	A	■■	Some A are not C		9	14
			No A are C		2	
			Some A are C		4	4
			Some C are A		2	1
			NVC			1
			Some C are not A		2	
O	I	■■	Some A are not C		9	15
			Some A are C		5	2
			NVC		5	3
O	E	■■	No A are C	10	15	9
			Some A are not C	5	2	
			NVC			10
			Some A are C	2	1	
			No A are (not) C	2	1	
O	O	■■	Some A are not C		14	13
			No A are C		1	
			NVC			6
			Some A are C		4	1

Table A2
Responses given by children, adolescents, and adults to syllogisms of figure BA-CB

	First premise					
A	A	□	ALL C ARE A	12	6	9
			SOME A ARE C		1	
			Some C are A			2
			All A are C	6	12	8
A	I	■■	All C are A	6	2	
			Some A are C	8	12	7
			Some C are A	3	4	12
			NVC			1
			All A are C	2	1	
A	E	□	NO C ARE A	5	3	10
			NO A ARE C	5	14	7
			All C are A	5	1	1
			All A are C	2	1	
			Nvc		1	2
A	O	■■	Some A are C		9	2
			Some C are not A		3	10
			No C are A		1	2
			No A are C		1	
			Some C are A			1
			NVC			3
			Some A are not C		6	2
I	A	□□	All C are A	1		
			Nvc			2
			SOME A ARE C	6	9	5
			SOME C ARE A	7	8	13
			All A are C	5	2	
I	I	■■	Some C are A	7	4	12
			Some A are C	12	10	7
			NVC		4	1
I	E	□□	No C are A	2		1
			No A are C	11	7	7
			Nvc		1	5
			SOME C NOT A		4	3
			Some C are A	5	5	4
			Some A are C		3	
I	O	■■	Some C not A		4	8
			Some A are C		7	1
			Some C are A		1	4
			No C are A			2
			No A are C		2	
			NVC		3	5
			Some A not C		3	
E	A	□□	No C are A		8	13
			No A are C		11	4
			Nvc			2
			SOME A ARE NOT C		1	
E	I	□□	No C are A		4	8
			No A are C		7	2
			Nvc			7
			SOME A ARE NOT C		5	2
			Some A are C		4	1
E	E	■■	No C are A		5	3
			No A are C		13	9
			NVC		1	8
E	O	■■	No C are A	10	3	8
			No A are C	3	8	2
			Some C are not A	1	1	
			NVC			7
			Some C are A	2	1	2
			Some A are C	3	5	
			Some A are not C	1	2	
O	A	■■	Some C are A		4	5
			Some A are C		4	1
			No A are C		3	1
			NVC			4
			Some C are not A		5	7
			Some A are not C		2	2
O	I	■■	Some C are A		2	3
			Some A are C		5	
			Some C are not A		2	5
			No C are A		1	
			No C are A		1	
			NVC		4	10
			Some A are not C		5	2
O	E	■■	No C are A			1
			No A are C		8	1
			Some C are not A		1	2
			NVC		2	12
			Some C are A		7	4
O	O	■■	Some C are not A		4	7
			NVC		2	6
			Some C are A		1	3
			Some A are C		6	
			Some A are not C		6	4

Table A3
Responses given by children, adolescents, and adults to syllogisms of figure AB-CB

	First premise					
A	A	■■	All A are C		16	8
A	A		Some A are C			1
A	A		NVC		2	10
A	I	■■	Some A are C		13	7
A	I		Some C are A		1	2
A	I		NVC		1	11
A	E	□	NO A ARE C	13	17	12
A	E		NO C ARE A	1		1
A	E		Nvc	2		7
A	O	□□	Some A are C	4	7	5
A	O		No A are C		1	
A	O		All C are not A	2		
A	O		Some C are A	1	1	1
A	O		Nvc	4	1	9
A	O		SOME A ARE NOT C	6	9	5
I	A	■■	All A are C	8	1	
I	A		Some C are A	4	10	6
I	A		Some A are C	3	6	2
I	A		NVC	3	3	12
I	I	■■	Some A are C	9	8	
I	I		Some C are A			1
I	I		NVC	7	11	13
I	E	□□	No A are C		11	6
I	E		No C are A		1	3
I	E		Nvc		2	8
I	E		SOME C ARE NOT A		2	2
I	E		Some A are C		3	
I	O	■■	Some A are C		5	2
I	O		Some C are A		2	
I	O		Some A are not C		6	3
I	O		No A are C		1	
I	O		NVC		5	13
I	O		Some C are not A		1	2
E	A	□	NO A ARE C	13	7	5
E	A		NO C ARE A	3	11	10
E	A		Nvc	2		5
E	I	□□	No A are C	3	5	2
E	I		No C are A	2	7	5
E	I		Nvc	3	4	7
E	I		SOME A ARE NOT C	3	3	3
E	I		Some A are C	6	1	3
E	E	■■	No A are C		8	2
E	E		NVC		10	18
E	O	■■	No A are C		5	1
E	O		No C are A		2	6
E	O		Some A are not C		7	1
E	O		NVC		1	12
E	O		Some A are C		3	
O	A	□□	All A are C	2	1	
O	A		Some C are A	3	2	3
O	A		Some A are C	2	4	1
O	A		No A are C		2	
O	A		All A are not C	8		
O	A		Nvc	2		4
O	A		SOME C ARE NOT A	2	4	7
O	A		Some A are not C		7	2
O	I	■■	Some A are C		4	1
O	I		Some C are A		1	
O	I		Some C are not A		3	
O	I		NVC		10	16
O	I		Some A are not C		2	2
O	E	■■	No A are C		8	5
O	E		Some A are not C		3	
O	E		Some C are not A		5	
O	E		NVC		2	13
O	E		Some C are A			2
O	O	■■	Some A are not C		7	4
O	O		No A are C		1	
O	O		NVC		8	15
O	O		Some A are C		3	1

Table A4
Responses given by children, adolescents, and adults to syllogisms of figure BA-BC

	First premise					
A	A	□	SOME A ARE C		1	3
A	A		Nvc	4	5	8
A	A		All A are C	12	14	9
A	A		All C are A	3		
A	I	□□	All A are C	1		
A	I		Nvc		1	1
A	I		SOME C ARE A	1	1	9
A	I		SOME A ARE C	14	15	8
A	E	□□	No A are C		15	10
A	E		No C are A		3	4
A	E		Nvc			3
A	E		SOME C ARE NOT A			1
A	E		Some A are C		2	
A	O	□□	All A are C		1	
A	O		Some C are A	2	2	1
A	O		Some A are C	7	9	1
A	O		All C are (not) A	3		1
A	O		Nvc			4
A	O		SOME C ARE NOT A	1	4	11
A	O		Some A are not C	4	3	1
I	A	□□	All C are A		1	
I	A		Nvc	1	1	1
I	A		SOME A ARE C	6	13	14
I	A		SOME C ARE A	3	2	5
I	A		All A are C	7	1	
I	I	■■	Some A are C	12	11	6
I	I		Some C are A	3		3
I	I		NVC	3	8	11
I	E	□□	No A are C	12	13	6
I	E		No C are A	3		1
I	E		Some A are not C		1	
I	E		Nvc	1	1	10
I	E		SOME C ARE NOT A		1	2
I	E		Some A are C	1	2	
I	E		Some C are A	3	1	1
I	O	■■	Some C are not A		2	3
I	O		Some A are C		13	2
I	O		Some C are A		1	1
I	O		NVC			13
I	O		Some A are not C		4	1
E	A	□□	No A are C		12	13
E	A		No C are A		2	2
E	A		Nvc		1	3
E	A		SOME A ARE NOT C			1
E	A		All A are C		3	1
E	I	□□	No A are C		5	1
E	I		No C are A		4	4
E	I		Nvc		1	9
E	I		SOME A ARE NOT C		1	4
E	I		Some A are C		7	2
E	E	■■	No A are C		10	4
E	E		No C are A		1	1
E	E		NVC		9	14
E	O	■■	No A are C		6	
E	O		No C are A		1	2
E	O		Some A are not C		1	2
E	O		NVC		2	12
E	O		Some C are not A			2
E	O		Some A are C		4	2
O	A	□□	Some A are C	4	4	2
O	A		Some C are A	2	1	
O	A		No A are C	7	3	
O	A		Nvc			3
O	A		SOME A ARE NOT C	3	7	14
O	A		Some C are not A	2	2	
O	I	■■	Some A are not C		8	9
O	I		Some C are A			1
O	I		Some A are C		6	1
O	I		NVC		6	9
O	E	■■	No A are C		7	2
O	E		No C are A		1	
O	E		Some A are not C		2	1
O	E		NVC		5	16
O	E		Some C are not A		2	
O	E		Some A are C		2	
O	O	■■	Some A are not C		9	5
O	O		No A are C		1	1
O	O		NVC		8	13
O	O		Some A are C		2	

References

Amidon, A. (1976). Children's understanding of sentences with contingent relations: why are temporal and conditional connectives so difficult? *Journal of Experimental Child Psychology, 22,* 423–437.

Baddeley, A. D. (1986). *Working memory.* Oxford: Oxford University Press.

Bar-Hillel, Y. (1967). Dictionaries and meaning rules. *Foundation of Language, 3,* 409–414.

Bara, B. G. (1995). *Cognitive Science: A developmental approach to the simulation of the mind.* Hove, UK: Lawrence Erlbaum Associates.

Bara, B. G., Bucciarelli, M., & Johnson-Laird, P. N. (1995). Development of syllogistic reasoning. *The American Journal of Psychology, 108* (2), 157–193.

Bara, B. G., Bucciarelli, M., Johnson-Laird, P. N., & Lombardo, V. (1994). Mental models in propositional reasoning. In A. Ram, & K. Eiselt (Eds.), *Proceedings of the Sixteenth Annual Conference of the Cognitive Science Society* (pp. 15–20). Hillsdale, NJ: Lawrence Erlbaum Associates.

Bara, B. G., Bucciarelli, M., & Lombardo, V. (2000). In favor of a unified model of deductive reasoning. In: J. A. Garcìa-Madruga, N. Carriedo, & M.J. Gonzalez-Labra (Eds.), *Mental models in reasoning* (pp. 6–81). Madrid: Universidad Nacional de Educacion a Distancia.

Barres, P. E., & Johnson-Laird, P. N. (1997). Why is it hard to imagine what is false? In M.G. Shafto, & P. Langley (Eds.), *Proceedings of the Nineteenth Annual Conference of the Cognitive Science Society* (p. 859). Mahwah, NJ: Lawrence Erlbaum Associates.

Bell, V., & Johnson-Laird, P. N. (1998). A model theory of modal reasoning. *Cognitive Science, 22,* 25–51.

Bloom, L., Lahey, M., Hood, L., Lifter, K., & Fiess, K. (1980). Complex sentences: acquisition of syntactic connectives and the semantic relations they encode. *Journal of Child Language, 7,* 235–261.

Bonatti, L. (1994). Why should we abandon the mental logic hypothesis? *Cognition, 50,* 17–39.

Bower, G. H., & Morrow, D. G. (1990). Mental models in narrative comprehension. *Science, 247,* 44–48.

Braine, M.D.S. (1978). On the relation between the natural logic of reasoning and standard logic. *Psychological Review, 85,* 1–21.

Braine, M. (1990). The 'natural logic' approach to reasoning. In: Overton, W.F. (Ed.), *Reasoning, Necessity, and Logic* (pp. 13–158). Hove, UK: Lawrence Erlbaum Associates.

Braine, M. (1998). Steps towards a mental predicate logic. In: Braine, M. D. S., & O'Brien, D. P. (Eds.), *Mental Logic* (pp. 27–331). Mahwah, NJ: Lawrence Erlbaum Associates.

Braine, M. D. S., & Rumain, B. (1981). Development of comprehension of 'Or': Evidence for a sequence of competencies. *Journal of Experimental Child Psychology, 31,* 46–70.

Braine, M. D. S., & Rumain, B. (1983). Logical reasoning. In: J. H. Flavell, & E. M. Markman (Eds.), *Carmichael's handbook of child psychology. Vol. III: Cognitive development* (4th ed., pp. 263–339). New York: Wiley.

Bucciarelli, M., & Johnson-Laird, P. N. (1999). Strategies in syllogistic reasoning. *Cognitive Science, 23,* 247–303.

Byrne, R. M. J. (1997). Cognitive processes in counterfactual thinking about what might have been. In D. L. Medin (Ed.), *The Psychology of Learning and Motivation, Advances Research and Theory, Vol. 37* (pp. 105–154). San Diego, CA: Academic.

Byrne, R. M. J., Handley, S. J., & Johnson-Laird, P. N. (1995). Reasoning from suppositions. *Quarterly Journal of Experimental Psychology, 48A,* 915–944.

Byrne, R. M. J., & Johnson-Laird, P. N. (1989). Spatial reasoning. *Journal of Memory and Language, 28,* 564–575.

Byrne, R. M. J., & McEleny, A., (2000). Counterfactual thinking about actions and failures to act. *Journal of Experimental Psychology: Learning Memory & Cognition, 26,* 1318–1331.

Chater, N., & Oaksford, M. (1999). The probability heuristics model of syllogistic reasoning. *Cognitive Psychology, 38,* 191–258.

Cheng, P. W., & Holyoak, K. J. (1985). Pragmatic reasoning schemas. *Cognitive Psychology, 17,* 391–416.

Cheng, P. W., Holyoak, K. J., Nisbett, R.., & Oliver, L. (1986). Pragmatic versus syntactic approaches to training deductive reasoning. *Cognitive Psychology, 18,* 293–328.

Clark, H. H. (1969). Linguistic processes in deductive reasoning. *Psychological Review, 76,* 387–404.

Cosmides, L., & Tooby, J. (1994). Evolutionary psychology and the generation of culture, Part II. Case study: A computational theory of social exchange. *Ethology and Sociobiology, 10,* 51–97

van Deemter, K., & Peters, S. (1996). *Semantic ambiguity and underspecification.* Stanford: CSLI publications.

DeSoto, L. B., London, M., & Handel, L. S. (1965). Social reasoning and spatial paralogic. *Journal of Personality and Social Psychology, 2,* 513–521.

Devetag, M. G., Legrenzi, P., & Warglien, M. (2000). Focusing strategies in reasoning about games. In: W. Schaeken, G. De Vooght, A. Vandierendonck, & G. d'Ydewalle (Eds.), *Deductive reasoning and strategies* (pp. 28–299). NJ, USA: Lawrence Erlbaum Associates.

Devlin, K. (1997). *Goodbye Descartes: The End of Logic and the Search for a New Cosmology of the Mind.* New York: Wiley.

Elio, R., & Pelletier, F. J. (1997). Belief change as propositional update. *Cognitive Science, 21,* 4, 419–460

Evans, J. St. B.T. (1989). *Bias in human reasoning. Causes and consequences.* East Sussex, U.K.: Lawrence Erlbaum Associates.

Evans, J. St. B.T. (1993). The mental model theory of conditional reasoning: critical appraisal and revision. *Cognition, 48,* 1–20.

Evans, J. St. B. T., Newsted, S. E., & Byrne, R. M. J. (1993). *Human reasoning: The psychology of deduction.* Hillsdale, NJ.: Lawrence Erlbaum Associates.

Evans, J. St. B. T., & Over, D. E. (1996). *Rationality and reasoning.* Hove, U.K.: Psychology Press.

Ford, M. (1994). Two modes of mental representation and problem solution in syllogistic reasoning. *Cognition, 54,* 1–71.

Garcìa-Madruga, J. A., Carriedo, N., & Gonzalez-Labra, M.J. (Eds.) (2000). *Mental models in reasoning.* Madrid: Universidad Nacional de Educacion a Distancia.

Garnham, A., & Oakhill, J. (Eds.). (1996). *Mental Models in Cognitive Science: Essays in Honour of Phil Johnson-Laird.* Hove, U.K.: Psychology Press.

Geminiani, G. C., Carassa, A., & Bara, B. G. (1996). Causality by contact. In: Garnham, A., & Oakhill, J. (Eds.), *Mental Models in Cognitive Science: Essays in Honour of Phil Johnson-Laird.* Hove, U.K.: Lawrence Erlbaum Associates.

Gigerenzer, G., & Hug, K. (1992). Domain-specific reasoning: Social contracts, cheating, and perspective change. *Cognition, 43,* 127–171.

Goldvarg, Y., & Johnson-Laird, P. N. (2000). Illusions in modal reasoning. *Memory & Cognition, 28,* 282–294.

Griggs, R. A., & Cox, J. R. (1982). The elusive thematic-materials effect in Wason's selection task. *Current Psychological Research & Reviews, 3,* 3–10.

Guyote, M. J., & Sternberg, R. J. (1981). A transitive-chain theory of syllogistic reasoning. *Cognitive Psychology, 13,* 461–525.

Hagert, G. (1984). Modeling mental models: Experiments in cognitive modeling of spatial reasoning. In T. O'Shea (Ed.), *Advances in artificial intelligence.* Amsterdam: North-Holland

Harman, G. (1986). *Change in view: Principles of reasoning.* Cambridge, MA: Bradford Books, MIT Press.

van der Henst, J. B. (1999). The mental model theory of spatial reasoning re-examined: the role of relevance in premise order. *British Journal of Psychology, 90*(1), 73–84.

Hunter, I. M. L. (1957). The solving of three-term series problems. *British Journal of Psychology, 48,* 286–298.

Huttenlocher, J. (1968). Constructing spatial images: a strategy in reasoning. *Psychological Review, 75,* 286–298.

Inhelder, B., & Piaget, J. (1958*). The growth of logical thinking from childhood to adolescence.* New York: Wiley.

Johansson, B. S., & Sjolin, B. (1975). Preschool children's understanding of the coordinators 'and' and 'or'. *Journal of Experimental Child Psychology, 19,* 233–240.

Johnson-Laird, P. N. (1983). *Mental models.* Cambridge, U.K.: Cambridge University Press.

Johnson-Laird, P. N. (1997). Rules and Illusions [A critical study of Rips's *The Psychology of Proof*]. *Minds and Machines, 7,* 387–407.

Johnson-Laird, P. N., & Bara, B. G. (1984). Syllogistic inference. *Cognition, 16,* 1–61.
Johnson-Laird, P. N., & Byrne, R. (1990). Meta-logical puzzles: Knights, knaves and Rips. *Cognition, 36,* 69–84.
Johnson-Laird, P. N., & Byrne, R. (1991). *Deduction.* Hillsdale, NJ.: Lawrence Erlbaum Associates.
Johnson-Laird, P. N., Byrne, R., & Schaeken, W. (1992). Propositional reasoning by model. *Psychological Review, 3,* 418–439.
Johnson-Laird, P. N., Byrne, R., & Tabossi, P. (1989). Reasoning by model: The case of multiple quantification. *Psychological Review, 96,* 658–673.
Johnson-Laird, P. N., Legrenzi, P., Girotto, V., Sonino-Legrenzi, M. S., & Caverni, J-P. (1999). Naive probability: A mental model theory of extensional reasoning. *Psychological Review, 106,* 62–88.
Johnson-Laird, P. N., Oakhill, J., & Bull, D. (1986). Children's syllogistic reasoning. *Quarterly Journal of Experimental Psychology, 38A,* 35–58.
Johnson-Laird, P. N., & Savary, F. (1996). Illusory inferences about probabilities. *Acta Psychologica, 93,* 69–90.
Johnson-Laird, P. N., Savary, F., & Bucciarelli, M. (2000). Strategies and tactics in reasoning. In: W. Schaeken, G. De Vooght, A. Vandierendonck, & G. d'Ydewalle (Eds.), *Deductive reasoning and strategies* (pp. 209–240). NJ, USA: Lawrence Erlbaum Associates.
Karmiloff-Smith, A. (1992). *Beyond modularity: A developmental perspective on cognitive science.* Cambridge, MA: MIT, Press.
Legrenzi, P., Girotto, V., & Johnson-Laird, P. N. (1993). Focusing in reasoning and decision making. *Cognition, 49,* 37–66.
Manktelow, K. I., Fairley, N., Kilparick, S. G., & Over, D. E. (2000). Pragmatics and strategies for practical reasoning. In: W. Schaeken, G. De Vooght, A. Vandierendonck, & G. d'Ydewalle (Eds.), *Deductive reasoning and strategies* (pp. 11–130). NJ, USA: Lawrence Erlbaum Associates.
Newell, A. (1990). *Unified theories of cognition.* Cambridge, MA: Harvard University Press.
Oakhill, J., Johnson-Laird, P. N., & Garnham, A. (1989). Believability and syllogistic reasoning. *Cognition, 31,* 117–140.
Oaksford, M., & Chater, N. (1994). A rational analysis of the Selection Task as optimal data selection. *Psychological Review, 101,* 608–631.
Piaget, J. (1953). *Logic and psychology.* Manchester: Manchester University Press.
Platt, R. D., & Griggs, R. A. (1993). Darwinian algorithms and the Wason selection task: A factorial analysis of social contract selection task problems. *Cognition, 48,* 163–192.
Polk, T. A., & Newell, A. (1995). Deduction as Verbal Reasoning. *Psychological Review, 102* (3), 533–566.
Pollard, P. (1981). The effect of thematic content on the 'Wason selection task'. *Current Psychological Researches, 1,* 21–29.
Pylyshyn, Z. W. (1973). What the mind's eye tells the mind's brain: a critique of mental imagery. *Psychological Bulletin, 80,* 1–24.
Rips, L. J. (1990). Reasoning. *Annual Review of Psychology, 41,* 321–353.
Rips, L. J. (1994). *The psychology of proof: Deductive reasoning in human thinking.* Cambridge, MA: MIT Press.
Rubin, E. (1921). *Visuell Wahrgenommene Figuren.* Copenhagen, Gyldendals.
Sacco, K., Bucciarelli, M., & Adenzato, M. (2001). Mental models and the meaning of connectives: A study on children, adolescents and adults. In J. D. Moore, & K. Stenning (Eds.), *Proceedings of the XXIII Conference of the Cognitive Science Society* (pp. 875–880). Edinburgh, Scotland: Lawrence Erlbaum Associates.
Schaeken, W., Johnson-Laird, P. N., & d'Ydewalle, G. (1996). Mental models and temporal reasoning. *Cognition, 60,* 205–234.
Sperber, D., Cara, F., & Girotto, V. (1995). Relevance theory explains the selection task. *Cognition, 57,* 31–95.
Staudenmayer, H., & Bourne, JR. L. E. (1977). Learning to interpret conditional sentences: a developmental study. *Developmental Psychology, vol. 13,* 6, 616–623.

Stenning, K., & Oberlander, J. (1995). A cognitive theory of graphical and linguistic reasoning: Logic and implementation. *Cognitive Science, 19* (1), 97–140.

Sternberg, R. J. (1985). *Beyond I.Q.: A triarchic theory of human intelligence.* New York: Cambridge University Press.

Vandierendonck, A., De Vooght, G., & Dierckx, V. (2000). Spatial and temporal content and working memory usage in linear syllogistic reasoning. In: J. A. Garcìa-Madruga, N. Carriedo, & M.J. Gonzalez-Labra (Eds.), *Mental models in reasoning* (pp. 189–199). Madrid: Universidad Nacional de Educacion a Distancia.