Neighborhoods of Successive Mental Models

Damien Raczy
Laboratoire de Psychologie Cognitive
Université Charles de Gaulle
Villeneuve d'Ascq, F-59650, FRANCE
raczy@univ-lille3.fr

Introduction

Here is sketched a formalization of mental model theory. It is shown that, in a reasoning task, two successive mental models differ from each other in one and only one token e.g. they are neighbors.

The formalization

Johnson-Laird and Byrne (1991) suggest that each entity, property and relation of the world should be represented by a corresponding token, property and relation. This is formalized by:

Definition *Basic sets of MMT:* Let O be a set of mental objects. Let M be a set of mental models, T a set of tokens, P a set of properties, V a set of values, R a set of references, such as $O = M \cup T \cup P \cup V \cup R$ and:

$$M = \{\{t_1, \ldots, t_n\} \subseteq T\}$$
$$T = \{\{p_1, \ldots, p_n\} \subseteq P\}$$
$$P = \{\{\langle v, r\rangle\} \subseteq V \times R\}$$

One of the basics of mental models computation can be summed up as following:

Principle *Three main operations*: Building mental models needs three operations:

- creating a new empty model: $\exists m \in M, m \leftarrow \varnothing$;
- add or subtract one token in a mental model: $m \leftarrow m \cup \{t\}$;
- combine two models together: $m_3 \leftarrow m_1 * m_2$.

Johnson-Laird and Byrne (1991) propose that the combining relies on iterative composing of tokens. So, in order to combine two models, it is necessary to compose tokens of these models, which is defined as:

Definition *Composition of tokens: t-composition* If $t_1 \cap t_2 \neq \varnothing$, the t-composition of two tokens t_1 and t_2 is $t_1 {}^\circ t_2 = t_1 \cup t_2$ else it is $t_1 {}^\circ t_2 = \varnothing$.

Definition *Combining of models:* The combining of two models m_1 and m_2 is:

$$m_1 * m_2 = \{m | \forall t \in m, \exists t_1, t_2 \in m_2, m_1, t = t_1 {}^\circ t_2\}$$

Neighborhoods

Combining of models is done step by step by the mean of adding or suppressing one token at a time. So, given m a mental model and $\pi = \{\ldots, p_i, \ldots\}$ a set of properties, the next mental model which can be built is $m \cup \{p_i\}$ or $m - \{p_i\}$. As a result, for all m_1, m_2 such that m_2 immediately follows m_1 in the reasoning, $|m_1 \otimes m_2| = 1$.[1] Adapting Freksa (1991), we propose:

Definition *Conceptual neighbors:* The mental models m_1 and m_2 are conceptual neighbors if and only if $|m_1 \otimes m_2| = 1$, which is noted $neigh(m_1, m_2)$

We informally define a conceptual neighborhood of mental models, or neighborhood, as a connected graph whose nodes define a set of mental models and edges are all possible 2-uples $\langle m_1, m_2\rangle$ such as $neigh(m_1, m_2)$.

This defines the graph of successive mental models that can be viewed as a conceptual neighborhood.

Empirical implication

Raczy & Anceaux (1996) and Raczy (1997) have validated this formalization in various formal deduction tasks: categorical, temporal, conditional, relational, spatial, hypothetico-deductive reasoning... These confirmed that subjects always build successive mental models which are neighbors.

References

Johnson-Laird P.N., Byrne R.M.J (1991) *Deduction*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Freksa, C. (1991). Conceptual neighborhood and its role in temporal and spatial reasoning. In *Proceedings of the international workshop on decision support systems and qualitative reasoning*. M. G. Singh and L. Travé-Massuyés, (pp. 181-187). Amsterdam, North-Holland.

Raczy D. (1997). Modèles mentaux et voisinages conceptuels: indétermination et contexte. Doctoral dissertation. Villeneuve d'ascq, France: Université de Lille 3.

Raczy D. & Anceaux F. (1996). Mental models, conceptual neighbourhoods and connexity. In *Proceedings of the International Conerence on Cognitive Systems*. Iran.

Johnson-Laird P.N., Byrne R.M.J. & Schacken W (1992). Propositional reasoning by model. *Psychological review*, *99*, 418-439.

1. Operator $\otimes$ is the symmetrical difference defined by $a \otimes b = a \cup b - a \cap b$.