In C. Delrieux and J. Legris (Eds.), (2003) Computer Modeling of Scientific Reasoning. Proceedings of the Third International Workshop on Computational Models of Scientific Reasoning and Applications, held at Buenos Aires, Argentina, September 14-15, 2003. Argentina: Ediuns, 189-198.

396

Semantic Computations of Truth Based on Associations already Learned

Patrick Suppes† and Jean-Yves Béziau‡

†CSLI - Stanford University, Stanford, CA, USA
‡Institute of Logic - Neuchâtel, Switzerland

Abstract. We present a new theory of how one computes truth based on associative networks.

*Keywords:*truth computation, associative networks, semantics.

1 Introduction: the problem of truth computation

In this paper we are trying to give an account of how one determines the truth or falsity of sentences like:

- Paris is the capital of France
- Paris is not the capital of France
- Rome is the capital of France
- Bangkok is not the capital of France
- Zürich is the capital of Switzerland

We want to describe the computations with the mental images of words and other things underlying such replies, taking in account in particular the time factor. Our theory should be able to explain the data gathered by experimentation, for example why it takes more time to give a negative answer (with false or/and with not) than a positive one, be it true or false. The general mechanism leading to a false or to a true reply should be the same, since there seems to have no differences in nature in the computation of the result.

Throughout this paper we refer to the mental images of words, and in some cases, visual maps of geographic regions. We mean nothing esoteric by "mental image". We use this term to refer in fact to brain-wave images, but the details of our views on brain-wave representations of words and sentences cannot be entered into here (see e.g. [8]). This is the reason we use the more general term "mental image".

2 Background of the theory

2.1 Philosophy and logic

Philosophers discuss at length various theories of truth (coherence theory, correspondence theory, etc.), problem of direct reference, sense and denotation and so on, but curiously are not able to give an account of how we perform truth-statements, and even less why we are able to perform them so quickly. Philosophers who claim that "Paris is the capital of France" is true because Paris is the capital of France are generally not interested to explain how we get from the fact to the assertion of truth of the corresponding sentence.

Logicians don't either solve these problems. If we want to describe how one answers a question like "Is 22+13 equal to 36?" it is certainly the wrong way to look at the logical foundations of arithmetics (Axioms of arithmetic), whether it is proof-theoretical or model-theoretical.

We answer to a question like "Is 22+13 equal to 36?" by using a series of small computational tricks, not by looking for a formal proof from a set of axioms or by finding a model in which the axioms are true and 22+13=36 is false.

In the case of a question like "Is Rome the capital of France?", it is even more doubtful that we are trying to deduce the truth or falsity of the sentence "Rome is the capital of France" from a set of axioms, or by drawing a truth-table.

From this point of view it is misleading to say that we are making a *deduction* to reach the conclusion that "Rome is the capital of France" is false, unless we strongly emphasized that deduction here does not reduce to the narrow meaning of deduction in formal logic. To avoid misunderstanding, it is better here to say that we are trying to describe how we *compute* the truth and falsity of a sentence like "Rome is the capital of France".

Logicians do not deal with this kind of problem, as stressed by Woods, they "have generally stopped short of trying to actually specify the truth conditions of the basic atomic propositions in their systems, dealing mainly with the specification of the meanings of complex expressions in terms of the meanings of elementary ones" ([9], p.220). According to Woods, researchers in artificial intelligence are trying to find an alternative solution where logicians failed. But do these people have a solution?

2.2 Artificial intelligence and computational linguistic

The *artificial intelligence paradox* is described as follows by Hölldobler commenting a paper by Shastri and Ajjanagadde, in which they propose a possible solution to this paradox: "the gap between the ability of humans to draw a variety of inferences effortlessly, spontaneously, and with remarkable efficiency on the one hand and the results about the complexity of reasoning reported by researchers in artificial intelligence on the other hand" ([5], p.463). This paradox shows very well that most of researches in the field of artificial intelligence don't solve the problem.

In order to compare our approach with the various approaches of AI researchers, or people working in computational linguistics, it is important to emphasize that generally it is not clear at all what they are trying to describe: reasoning, processing of language, etc. Several isssues are mixed and nobody is dealing specifically with the truth problem.

In computational lingustics people are more interested in syntax: formal grammars, parsing of sentences and so on. In AI, people are more oriented to wards semantics, for example they have developed "semantic networks". However the status of such networks is not clear from the viewpoint of the syntax/semantics distinction as show ambiguous expressions used by AI researchers like "semantics for semantic networks". Woods remarks that "the question of what (semantic) networks have to do with semantics is one which takes some answering" ([9], p.218). Distinctions between inference, truth and meaning are not clear at all in semantics networks which are a kind of mix of everything. Anyway it seems that the orientation taken by AI researchers is better than the one taken by computational linguists, because with semantics networks they are trying to find a shorter path without going into syntax and logical representation of natural language.

In a recent book on computational semantics, the authors, P.Blackburn and J.Bos, say:

> The book is devoted to introducing techniques for tackling the following two questions:
> 1. How can we automate the process of associating semantic representations with expressions of natural language?
> 2. How can we use logical representations of natural language expressions to automate the process of drawing inferences? ([1], p.iii)

Their idea is to find some algorithms to translate natural language into the language of first-order logic supposed to represent the meaning of natural language sentences and then to find some algorithms to describe inferences performed with these first-order translations. The two steps seem wrong for our purpose: it is doubtful that our brain use first-order logic to compute truth, and truth of atomic sentences is not based on logical inference rules.

Both AI researchers and computational linguists have been overinfluenced by logic. They don't deal directly with the problem of simple truth statements like "Paris is the capital of France". They include this problem in a general theory of reasoning based on logical inference, semantical or syntaxical.

2.3 Associations

Our common point with AI researchers is our emphasis on *associations* (sometimes in AI, "semantics networks" are also called "associative networks"). When answering a question like "Is Paris capital of France?" we are using notions which are associated with the input, like Eiffel Tower with Paris, country with capital, wine and cheese with France and so on. Our purpose here is to try to explain the mechanism of such associations, in connection with the question of truth and falsity (in this point we differ with AI reseachers who are concerned with broader problems).

Our description should not depend crucially on words-language, even though we are working with words-language examples, since we think that this mechanism has a common root in processes involving any signs-language, animal behaviour, stimulus-response phenomena in general.

For us an associative network is a set of objects with links between them. One central question is how is organized an associative network. An interesting proposal of global organization of lexicon in English based in psycholonguistic consideration has been provided by the *Wordnet* project, the organization is based on three pairs of relationship: antinomy-synonymy, hypernymy-hyponymy, meronymy-holonomy (See [3], [6]). However this classification presents several serious limitations, in particular for the analysis of the truth problem.

We will not present here any general theory of organization or hierarchization of associative networks, we will just focus our attention on how truth fits into associative networks. For us truth is here no mysterious entity, but just a word representing a mental image in the associative network, like "Paris" or "capital". Our main task is to explain, given an input like "Is Paris the Capital of France?", what happens in the associative network. Our idea is that the mental images represented by "truth" or "false" become linked with the mental image represented by "Paris is the capital of France", on the basis of some already existing links. We suppose here that these links are fixed, that they correspond to associations already learned, but of course in reality the links have to be considered dynamically.

3 The theory

3.1 Informal explanation

We consider a network which goes through different states. The nodes of the networks are mental images of words, visual images, etc. Among the mental images of words of the networks we have *true*, *false* (alethic words) and *not* (the only logical word) that may or may not be part of a sentence. Words which are neither alethic or logical are called proper words.

At the initial state of the network, we have some links which represent the associations already learned. We say that these links are *quiescent*. These links will later on become *activated* or not depending on the situation. We consider therefore that there are only two states for a link: quiescent or activated. This is of course a simplification. The nodes can also be activated or not. At the initial state, no mental image of a word is activated.

In the statement of the axioms, we use only qualitative probabilities that are high, i.e., close to 1, or small, i.e., close to 0. It will be possible to estimate numerically these probabilities from the error rates of individuals processing in experiments our sample geography sentences. From preliminary data we know the qualitative probabilities used in the axioms will be confirmed but the estimate of actual numerical values will vary from one individual to another and also from one experiment to another depending on the relative difficulty of the geography sentences presented.

<u>Axioms for the initial state</u>
Axiom I1 No mental image of a word is activated.
Axiom I2 All the links are quiescent.

Axiom I3 There are no links between mental images of alethic words ("true", "false") and mental images of proper words.
Axiom I4 There are quiescent links between the mental images of "false" and "not", and between those of "true" and "not".

Axioms for activation

Axiom A1 Given an input sentence S, the probability of the mental image of a word of S already in the network to be activated is high.
Axiom A2 Given an input sentence S, the probability of the mental image a word of S not already in the network to be included in the network and be activated is high.
Axiom A3 Given an input sentence S, the probability of two or more mental images of words of S not being included or not being activated in the network is small.
Axiom A4 If the mental images of two words are activated and there is a quiescent link between them, then with high probability this link becomes activated.

Axioms for true

Axiom T1 If an activated link corresponding to the sentence S coincides with a link given by the input, then with high probability an hyperlink is activated between true and S.

Axioms for false

Axiom F1 If the mental images of "true" and "not" are activated, then with high probability the mental images of "false" is activated.
Axiom F2 If there is an activated hyperlink between the mental image of "true" and the link corresponding to the mental image of the sentence S and the mental image of "not" is activated, then with high probability a hyperlink is activated between the mental image of "false" and the link corresponding to the mental image of *notS*.

Axioms for spreading activation

Axiom S If after the time t_n, no hyperlink is activated with the mental images of "true" and "false", then with high probability neighborhood nodes of activated nodes are activated at the time t_{n+1}.

Axioms for maps

Axiom M1 If the mental image of a map is activated and the mental images of two words of the map are activated, if they are close in the map, then with high probability a link is activated between them in the network.
Axiom M2 If the mental image of a map is activated and the mental images of words of the map are activated, then mental images of words which are close to them in the map and are in the network are activated with high probability.

3.2 Examples

For all our examples we have the same initial state of the network, described as follows ($\sim$ is used to represent links, $\approx$ represents activated links and underlining means that the mental image of the word is activated):

$Paris \sim capital, Rome \sim capital$
$Paris \sim city, Rome \sim city, Zurich \sim city$
$France \sim country, Italy \sim country, Switzerland \sim country$
$France \sim capital, Italy \sim capital, Switzerland \sim capital$
$Paris \sim France, Rome \sim Italy, Zurich \sim Switzerland$
$capital \sim 1-1, capital \sim city$
$Paris \sim EiffelTower, Paris \sim wine, Rome \sim wine$
$true \sim not, false \sim not$

Example 1
Input: Paris is the capital of France
Answer: true

State 1
$\underline{Paris} \approx \underline{capital}, Rome \sim \underline{capital}$
$\underline{Paris} \sim city, Rome \sim city, Zurich \sim city$
$\underline{France} \sim country, Italy \sim country, Switzerland \sim country$
$\underline{France} \approx \underline{capital}, Italy \sim capital, Switzerland \sim \underline{capital}$
$\underline{Paris} \approx \underline{France}, Rome \sim Italy, Zurich \sim Switzerland$
$\underline{capital} \sim 1-1, \underline{capital} \sim city$
$\underline{Paris} \sim EiffelTower, \underline{Paris} \sim wine, Rome \sim wine$
$true \sim not, false \sim not$

State 2
$\underline{Paris} \approx \underline{capital}, Rome \sim \underline{capital}$
$\underline{Paris} \sim city, Rome \sim city, Zurich \sim city$
$\underline{France} \sim country, Italy \sim country, Switzerland \sim country$
$\underline{France} \approx \underline{capital}, Italy \sim capital, Switzerland \sim \underline{capital}$
$\underline{Paris} \approx \underline{France}, Rome \sim Italy, Zurich \sim Switzerland$
$\underline{capital} \sim 1-1, \underline{capital} \sim city$
$\underline{Paris} \sim EiffelTower, \underline{Paris} \sim wine, Rome \sim wine$
$\underline{true} \sim not, false \sim not$
$\underline{true} \approx \underline{Paris/Capital/France}$

Example 2
Input: Paris is not the capital of France
Answer: false

State 1
$\underline{Paris} \approx \underline{capital}, Rome \sim \underline{capital}$
$\underline{Paris} \sim city, Rome \sim city, Zurich \sim city$
$\underline{France} \sim country, Italy \sim country, Switzerland \sim country$
$\underline{France} \approx \underline{capital}, Italy \sim capital, Switzerland \sim \underline{capital}$
$\underline{Paris} \approx \underline{France}, Rome \sim Italy, Zurich \sim Switzerland$
$\underline{capital} \sim 1-1, \underline{capital} \sim city$
$\underline{Paris} \sim EiffelTower, \underline{Paris} \sim wine, Rome \sim wine$
$true \sim \underline{not}, false \sim \underline{not}$

State 2
$\underline{Paris} \approx \underline{capital}, Rome \sim \underline{capital}$
$\underline{Paris} \sim city, Rome \sim city, Zurich \sim city$
$\underline{France} \sim country, Italy \sim country, Switzerland \sim country$

$\underline{France} \approx \underline{capital},\ Italy \sim \underline{capital},\ Switzerland \sim \underline{capital}$
$\underline{Paris} \approx \underline{France},\ Rome \sim Italy,\ Zurich \sim Switzerland$
$\underline{capital} \sim 1-1,\ \underline{capital} \sim city$
$\underline{Paris} \sim EiffelTower,\ \underline{Paris} \sim wine,\ Rome \sim wine$
$\underline{true} \approx \underline{not},\ false \sim \underline{not}$
$\underline{true} \approx Paris/Capital/France$

State 3
$\underline{Paris} \approx \underline{capital},\ Rome \sim \underline{capital}$
$\underline{Paris} \sim city,\ Rome \sim city,\ Zurich \sim city$
$\underline{France} \sim country,\ Italy \sim country,\ Switzerland \sim country$
$\underline{France} \approx \underline{capital},\ Italy \sim \underline{capital},\ Switzerland \sim \underline{capital}$
$\underline{Paris} \approx \underline{France},\ Rome \sim Italy,\ Zurich \sim Switzerland$
$\underline{capital} \sim 1-1,\ \underline{capital} \sim city$
$\underline{Paris} \sim EiffelTower,\ \underline{Paris} \sim wine,\ Rome \sim wine$
$\underline{true} \approx \underline{not},\ \underline{false} \approx \underline{not}$
$\underline{true} \approx \underline{Paris/Capital/France}$
$\underline{false} \approx \underline{Paris/not/CapitalFrance}$

Example 3
Input: Bangkok is not the capital of France
Answer: true

Bangkok is not in the network, that means that the subject knows nothing about this city, but he arrives at the right conclusion due to the fact that he knows that Paris is the capital of France and that a country has only one capital (condition expressed by the mental image represented by "1-1").

State 1
$Paris \sim \underline{capital},\ Rome \sim \underline{capital}$
$Paris \sim city,\ Rome \sim city,\ Zurich \sim city$
$\underline{France} \sim country,\ Italy \sim country,\ Switzerland \sim country$
$\underline{France} \approx \underline{capital},\ Italy \sim \underline{capital},\ Switzerland \sim \underline{capital}$
$Paris \sim \underline{France},\ Rome \sim Italy,\ Zurich \sim Switzerland$
$\underline{capital} \sim 1-1,\ \underline{capital} \sim city$
$Paris \sim EiffelTower,\ Paris \sim wine,\ Rome \sim wine$
$true \sim not,\ false \sim \underline{not}$
$\underline{Bangkok}$

State 2
$\underline{Paris} \approx \underline{capital},\ \underline{Rome} \approx \underline{capital}$
$\underline{Paris} \approx \underline{city},\ \underline{Rome} \approx \underline{city},\ Zurich \sim \underline{city}$
$\underline{France} \approx \underline{country},\ Italy \sim \underline{country},\ Switzerland \sim \underline{country}$
$\underline{France} \approx \underline{capital},\ Italy \sim \underline{capital},\ Switzerland \sim \underline{capital}$
$\underline{Paris} \approx \underline{France},\ \underline{Rome} \sim Italy,\ Zurich \sim Switzerland$
$\underline{capital} \approx \underline{1-1},\ \underline{capital} \approx \underline{city}$
$\underline{Paris} \sim EiffelTower,\ \underline{Paris} \sim wine,\ \underline{Rome} \sim wine$
$true \sim not,\ false \sim \underline{not}$
$\underline{Bangkok}$

State 3
$\underline{Paris} \approx \underline{capital},\ \underline{Rome} \approx \underline{capital}$
$\underline{Paris} \approx \underline{city},\ \underline{Rome} \approx \underline{city},\ Zurich \sim \underline{city}$

$\underline{France} \approx \underline{country}, Italy \sim \underline{country}, Switzerland \sim \underline{country}$
$\underline{France} \approx \underline{capital}, Italy \sim \underline{capital}, Switzerland \sim \underline{capital}$
$\underline{Paris} \approx \underline{France}, \underline{Rome} \sim Italy, Zurich \sim Switzerland$
$\underline{capital} \approx \underline{1-1}, \underline{capital} \approx \underline{city}$
$\underline{Paris} \sim EiffelTower, \underline{Paris} \sim wine, \underline{Rome} \sim wine$
$\underline{true} \approx \underline{not}, \underline{false} \approx \underline{not}$
$\underline{Bangkok\ true} \approx \underline{Paris/capital/France}$
$\underline{false} \approx \underline{Bangkok/capital/France}$
$\underline{true} \approx \underline{Bangkok/not/capital/France}$

Example 4
Input: Zürich is the capital Switzerland
Answer: true

In this example, the subject first mistinguishes Switzerland with the Netherlands (confusion of audition), then activates a mental image of a map according to which Zürich is in Switzerland and finally concludes incorrectly that it is the capital of Switzerland (a common mistake).

State 1
$Paris \sim \underline{capital}, Rome \sim \underline{capital}$
$Paris \sim \underline{city}, Rome \sim \underline{city}, \underline{Zurich} \sim city$
$France \sim country, Italy \sim country, Switzerland \sim country$
$France \sim \underline{capital}, Italy \sim \underline{capital}, Switzerland \sim \underline{capital}$
$Paris \sim \underline{France}, Rome \sim Italy, \underline{Zurich} \sim Switzerland$
$\underline{capital} \sim 1-1, \underline{capital} \sim city\ Paris \sim EiffelTower$
$\underline{Paris} \sim wine, \underline{Rome} \sim wine$
$true \sim not, false \sim not$
$\underline{theNetherlands}$

State 2
$Paris \sim \underline{capital}, Rome \sim \underline{capital}$
$Paris \sim \underline{city}, Rome \sim \underline{city}, \underline{Zurich} \sim city$
$France \sim country, Italy \sim country, Switzerland \sim country$
$France \sim \underline{capital}, Italy \sim \underline{capital}, Switzerland \sim \underline{capital}$
$Paris \sim \underline{France}, Rome \sim Italy, \underline{Zurich} \sim Switzerland$
$\underline{capital} \sim 1-1, \underline{capital} \sim city\ Paris \sim EiffelTower$
$\underline{Paris} \sim wine, \underline{Rome} \sim wine$
$true \sim not, false \sim not$
$\underline{map} \approx \underline{Zurich}, \underline{map} \approx \underline{capital}, \underline{map} \sim Switzerland$
$\underline{theNetherlands}$

State 3
$Paris \sim \underline{capital}, Rome \sim \underline{capital}$
$Paris \sim \underline{city}, Rome \sim \underline{city}, \underline{Zurich} \sim city$
$France \sim country, Italy \sim country, Switzerland \sim country$
$France \sim \underline{capital}, Italy \sim \underline{capital}, \underline{Switzerland} \approx \underline{capital}$
$Paris \sim \underline{France}, Rome \sim Italy, \underline{Zurich} \approx \underline{Switzerland}$
$\underline{capital} \sim 1-1, \underline{capital} \sim city\ Paris \sim EiffelTower$
$\underline{Paris} \sim wine, \underline{Rome} \sim wine$
$true \sim not, false \sim not$
$\underline{map} \approx \underline{Zurich}, \underline{map} \approx \underline{capital}, \underline{map} \approx \underline{Switzerland}$
$\underline{theNetherlands}$

State 4

$Paris \sim \underline{capital}$, $Rome \sim \underline{capital}$
$Paris \sim city$, $Rome \sim city$, $\underline{Zurich} \sim city$
$France \sim country$, $Italy \sim country$, $Switzerland \sim country$
$France \sim \underline{capital}$, $Italy \sim \underline{capital}$, $\underline{Switzerland} \approx \underline{capital}$
$Paris \sim France$, $Rome \sim Italy$, $\underline{Zurich} \approx \underline{Switzerland}$
$\underline{capital} \sim 1-1$, $\underline{capital} \sim city$ $Paris \sim EiffelTower$
$Paris \sim wine$, $Rome \sim wine$
$\underline{true} \sim not$, $false \sim not$
$map \approx \underline{Zurich}$, $map \approx \underline{capital}$, $map \approx \underline{Switzerland}$
$\underline{theNetherlands}$
$\underline{true} \approx \underline{Zurich/capital/Switzerland}$

4 Philosophical consequences of the theory: the distinctions between analytic and synthetic, a priori and a posteriori revisited

From our point of view, there are no basic differences between the truth of "2+2=4" and "Paris is the capital of France": both result of computation on associative networks. It seems therefore difficult to say that the truth of "2+2=4" is analytic or a priori by opposition to the truth of "Paris is the capital of France" which would be synthetic or a posteriori.

We can however draw some distinctions between the computation of truth of different sentences. An important difference about the computation of different sentences is the time necessary to perform the computation. But can we say that the truth of sentences whose truth computation requires less time is analytic or a priori rather than synthetic or a fortiori? Then what is the time barrier which will draw the line? From this point of view a negative or false sentence will be less analytic than a true sentence.

It is better to focus on the level of activation necessary to compute the truth of a sentence. If no mental images of proper words other than the words of the sentence are necessary to compute the truth of it, then we can say that the truth or falsity of the sentence is analytic, by opposition we will say that the truth or falsity is synthetic if it requires the activation of mental images of other words not in the sentence. Example 1 is an analytic truth and Example 3 is a synthetic truth.

We can distinguish the pair analytic/ synthetic from the pair a priori/ a posteriori by saying that the truth of a sentence is a priori if its computation does not require something else outside of the network and the input. Example 4 deals with an a posteriori truth, since a mental image of a map not originally in the network is used.

From this point of view, there are synthetic a priori truths, for example "Bangkok is not the capital of France" is a synthetic a priori truth by opposition to "Paris is the capital of France" which is an a priori analytical truth (relatively to the given associative network).

We don't see therefore analytical truths as tautological truths of the type "The capital of France is the capital of France". In fact if we ask the question "Is the capital of France the capital of France?" to an ordinary man, he may have some difficulty to understand, and takes more time to give a positive answer to this question than to a question like "Is Paris is the capital of France?", because the question may sound to him as a nonsense. Much the same can be said about less tautological questions like "Is the capital of France in France?".

References

1. P.Blackburn and J.Bos. *Representation and inference for natural language (A first course in computational semantics).* http:// www. comsem.org, 1999.
2. L.Buchanan, C.Westbury and C.Burgess. "Characterizing semantic space: neighborhood effects in word recognition". *Psychonomic Bulletin and Review*, **8**, pp.531–544, 2001.
3. C.Felbaum. *Wordnet, an electronic lexical database for English.* MIT Press, Cambridge,1998.
4. F.Lehman. "Semantic networks". In *Semantics networks in artificial intelligence.* F.Lehman (ed), Pergamon Press, pp.1–50, 1992.
5. L.Shastri and V.Ajjanagadde. "From simple associations to systematic reasoning: a connectionist representation of rules, variables and dynamic bindings using temporal synchrony". *Behavioral and Brain Sciences*, **16**, pp.417–494, 1993.
6. M.Sigman and G.A.Cecchi. "Global organization of the lexicon". To appear in *Proceedings of the National Academy of Sciences*, 2002.
7. P.Suppes. "Stimulus-response theory of finite automata". *Journal of Mathematical Psychology*, **6**, pp.327–355, 1969.
8. P.Suppes, B.Han, J.Epelboim and Z.L.Lu. "Invariance between subjects of brainwave representations of language". *Proceedings of the National Academy of Sciences*, **96**, pp.12953–12958, 1999.
9. W.A.Woods. "What's in a link? (Foundations for Semantic Networks)". in *Readings in knowledge representation*, R.J.Brachman and H.J.Levesque (eds), Morgan Kaufmann, Los Altos, pp.217–241, 1985.