WHERE QUESTIONS

Benny Shanon
The Hebrew University of Jerusalem

Consider question (1), and the answers to it, (2)-(4):

(1) Where is the Empire State Building?

(2) In New York.

(3) In the U.S.A.

(4) On 34th Street and 3rd Avenue.

When (1) is posed in California (2) is the appropriate answer to it. This is the case even though (3) and (4) are also true characterizations of the location of the Empire State Building. The pattern of appropriateness alters, however, when the locale where the question presented changes. Thus, when (1) is asked in Israel, (3) is the appropriate answer, whereas when it is asked in Manhattan, (4) is the answer that should be given.

The foregoing observations, originally made by Rumelhart (1974) and by Norman (1973), suggest the following. First, it is not enough for answers to questions to be (semantically) true, they have to be (pragmatically) appropriate as well. Second, appropriateness is not solely determined by the content of the particular propositions in question, but also by the identity of the participants in the particular conversational situation and their locale. In other words, for a person--or for a machine, for that matter--to answer questions, it is not enough to survey one's memory and retrieve information pertaining to the query posed, rather--a selection algorithm has to be used so that an appropriate response would be given. The specification of such a selection algorithm is the topic of the present investigation.

The following discussion is based on what is known as the Room Theory: the original, albeit preliminary, model proposed by Rumelhart(1974) in order to account for his insightful observations. I try to examine the psychological validity of this model, and to propose amendments and extensions to it on the basis of empirical data.

The Room Theory "posits the existence of a psychological room relative to which distances are reckoned. The room corresponds to the smallest geographical region that encompasses both the reference location of the conversants and the location of the places in question". When answering where-questions "the rule is to find the smallest room which just includes the reference location and the answer location. The appropriate answer is the next smallest geographical unit which contains the location in question, but excludes the reference location". (Rumelhart, 1974). The answers generated by this algorithm, note, constitute the placing of the item questioned in a room which is larger than it; henceforth answers of this type will be called vertical.

In order to examine the Room Theory, questions regarding places in the world as well as objects in a (concrete) room were presented to several subject populations: college students in Israel and the U.S., American children of three age groups, and aphasic patients. The present report concentrates on the adult data, and only cursory remarks will be made on the answers furnished by the other populations. First, I will discuss answers solicited by an open questionnaire, in which subjects were asked to give one answer to the questions posed to them; later, answers solicited by closed questionnaires will be discussed.

First, it should be noted that by and large the answers given by subjects were the ones predicted by the Room Theory. The correspondence between the data and the Theory is on two counts. First, there is the seemingly trivial observation that answers to different questions are given on different levels. Specifically, there is a correlation between the level of the object which is queried and the room on which the respective answer is given. Second, and less trivial, is the observation that answers vary not only with the questions, but also with the spatial relationship which holds between the object of the question and the participants in the conversation. Several loci in the data are indicative of this last pattern. First none of the Americans indicated that the Empire State Building was "in the U.S.", but a third of the Israelis did so; further, some of the Americans, but none of the Israelis, indicated that the building was "in Manhattan". Second, asked about New York City, almost all Israelis, but none of the Americans, answered on the country level. Further, the distribution of the answer patterns furnished by the members of each group changed according to whether the queried city was their own, or close/distant from it. Finally, children's answers to questions about objects also vary with how distant the object is.

Above, however, I have qualified the correspondence between the data and the theory; this qualification should now be clarified. I don't think it is meaningful to judge the validity of a model like the Room Theory by examining the percentage of cases in which its predictions hold. Such a percentage may reflect the structure of the domain (questions) under investigation, and it need not be indicative of the adequacy of the model as such. The term "by and large" is, however, of qualitative significance. It indicates that unless other factors or reasons are operative, answers to where questions do, indeed, follow the Room Algorithm. The detection of these "other factors and reasons", their classification and the characterization of the answer types that correspond to them is the main theme of this discussion. Following, then, are the answer patterns which do not conform with the Room Theory.

First, consider questions about landmarks in the towns in which the conversation took place. Most of the answers which involved vertical placement were given on the level of the town itself, i.e. on a level which is higher than the one predicted by the Room Theory. The other answers were not vertical, but rather horizontal:the object questioned was related to another object similar to it. In other words, either the level specified by the Room Theory was changed, or the type of answer (i.e. the generation algorithm itself) was altered. These deviant answers are viewed as two alternative solutions to the problem of the floor effect. Specifically, as one goes down the place hierarchy, the specification of rooms between the target and the least common room is cumbersome; indeed, there might not be simple names by which reference to these rooms may be made. Subjects solve this problem either by staying on the level of the least common room or by shifting to the horizontal strategy.

The same problem is noted with the ceiling effect, namely, with questions regarding objects which are very high on the place hierarchy: continents for adults, countries for children and aphasic patients. The answers in these cases were varied, a feature which attests the algorithmic difficulty associated with them. Only a minority of the answers conformed with the Room Theory and most answers were horizontal. Other answer types were: vacuous, in which a vertical answer was given on too high a level (e.g. "in the world"), featural, in which

a description, rather than a specification of the locale, was given (e.g. "it is a continent"), or tautological (e.g. "Japan is in Japan"). The different answer types, we shall say, are the products of different alternative answer generation algorithms. The numerical distribution of these answers suggest that the order of preference for the application of the algorithms as the one noted above.

There were also cases in which subjects gave answers on a level lower than the one predicted by the Room Theory. Thus, half the Israelis placed the Empire State Building "in New York", and not "in the U.S." Similarly, all the Americans asked about the Eiffel Tower answered "in Paris", and not "in France". These patterns are attributed to prominence. Prominent objects are ones which gain a higher rank in the place hierarchy than would be attributed to them on semantic classificatory grounds alone. As a consequence, these objects are placed in a room which is more specific than the one predicted by the Room Theory. For instance, New York City is not conceived of by non-Americans as just another American city; it gains an autonomy of its own and is conceived of as independent of the country in which it is located. The prominence effect suggests that rather than interpreting the room-hierarchy in a concrete fashion (i.e. as isomorphic to the spatial relations which hold in the physical world), one should view it as an abstract conceptual representation. In this representation, objects are associated with tags: usually, objects which are actually contained in objects of order n are assigned a tag of order $n+1$, but prominent objects are assigned tags of the same order as the objects which actually contain them. Thus, if the Empire State Building is tagged $n+1$ both New York and the U.S. are tagged n, for the Israelis the least common room (order $n-1$) is the northern hemisphere, and the answer is given on the level of the two rooms of order n. Thus, the seemingly unexpected answers associated with prominent objects are due to the modified abstract representations, not to a change in the(vertical) algorithm proper.

The salience effect is similar, but distinct. Objects which are close to ones which stand in a particular relation to the respondent (i.e. physically close, emotionally dear, or belonging to the subject) are not placed in a room but receive horizontal answers instead. For example, all the Israelis answered that Lebanon was "north of Israel", and not that it was "in the Mid-East". Similarly, all the Americans (and half of the Israelis) placed Canada in relation to the U.S. Unlike the prominence effect, the salience effect does affect the answer generation algorithm itself, and it bears on individual or cultural differences, not on general semantic considerations.[4] Specifically, items which are special to the speaker are tagged in the representation as marked, and this triggers a shift from the vertical to the horizontal algorithm.

All questions considered so far involved one configuration: the two conversants and the target were physically distinct, and together they could be contained in one common room. This, however, is not the only possible configuration. Other configurations, are possible as well: (a) The conversants and the target may coincide in place, as in the question "Where are we now?". (b) The conversants may be contained in the target, as in the question "Where is Israel?" when posed in Jerusalem. (c) The conversants may be in different places, as in phone conversations.

Strictly speaking, the Room Algorithm does not apply to these configurations. Thus, in (b) the least common room is one level above that of the target, but on what level would the answer be? The Room Algorithm would either return the respondent to the place queried or else require detailed and perhaps cumbersome classifications. Neither option is taken. All the answers to the questions noted were given on the room immediately above the target. In (a) a least common room may not be circumscribed in the manner outlined by the Room Algorithm, whereas in (c) a distinction between the speaker and the hearer has to be introduced. All these cases suggest that the different configurations do invoke different generation algorithms. Hence, an appraisal of the configuration is necessary prior to the application of the answer-generation algorithm proper.

So far, the discussion was topological, considering only the spatial configuration holding between the conversants and the object questioned. The respondent's knowledge of the world was not taken into account. In order to prove the psychological validity of an answer generation algorithm it is crucial to demonstrate that the answer given is chosen from a class of several feasible answers, and is not the only one possible due to a limited data base. This was the purpose of the closed questionnaires. Two such questionnaires were administered: first, subjects were asked to choose the best of several answers given to them; then they were asked to mark all the answers they deemed true. Three points were of interest. First, the answers given in the first two conditions were not necessarily the most specified ones marked in the third. Second, there were answers in the multiple option condition which were evidently true and commonly known but which were nonetheless not marked by subjects. These answers included reversed prominence (i.e. the relation of a prominent object to a less prominent one), featural answers and ones which were too high on the place hierarchy. Third, an "I don't know" answer on the open questionnaire did not necessarily imply a no-answer in the other conditions. In other words, this answer does not signify complete ignorance, but rather an appreciation on the part of the subject that he cannot furnish the answer he deems appropriate. Together, the three points indicate that there is indeed a psychological process of answer-generation which does not amount to the specification of the most detailed information one has regarding the object in question.

Still another aspect which has to be considered is the speaker's intention when he poses a question. A study of this aspect is just on its way now and at this point, I have to limit myself only to a methodological discussion. Evidently, the process of question-answering requires an appraisal of intention (cf. Lehnert, 1978), one which involves the evaluation of various contextual, personal and sociological factors. In order to make research feasible, as well as constructive, a factorization of the domain of question-answering, I believe, is needed. In this regard the topological, knowledge and intention aspects were noted. The original Room Theory is an attempt to define the topological aspect. The present study shows that even for this aspect this Theory is not sufficient. The present discussion suggests that an extended topological theory should consist of the following components:

1. Semantic and episodic representations, which are not isomorphic to the physically (logically) defined room-hierarchy.

2. Determinants of configurations and problematic cases (floor, ceiling).

3. A set of ordered answer-generation algorithms: vertical placement (the algorithm proposed by the Room Theory), horizontal relation, featural description and non-informative (vacuous, tautological).

Definitely, the topological consideration is not sufficient for the characterization of how people answer where questions. Future investigations should extend the research and also include considerations of knowledge and intention. At this juncture, however, we can note

that it is not possible to reduce question answering to knowledge alone, and that some formal selection algorithms have to be postulated. The formal study of such algorithms is of relevance to the study of both natural and artificial intelligence.

References

Lehnert, W.G. The process of question answering. Lawrence Erlbaum and Associates, Hillsdale, N.J., 1978.

Norman, D. Memory, knowledge and the answering of questions. In R. Solso (Ed.) Contemporary Issues in Cognitive Psychology, Washington, D.C.: E. Winston, 1973.

Rumelhart, D. The Room Theory. Unpublished manuscript, The University of California at San Diego, 1974.