The Real Estate Agent—
Modeling Users By Uncertain Reasoning

Katherina Morik
Claus-Rainer Rollinger

Technische Universität Berlin Projekt KIT, Institut für Angewandte Informatik, Franklinstrasse 28/29, 1000 Berlin 10

Abstract
Two topics are treated here. First, we present a user model patterned after the stereotype approach (Rich, 1979). This model surpasses Rich’s model with respect to its greater flexibility in the construction of user profiles, and its treatment of positive and negative arguments. Second, we present an inference machine. This machine treats uncertain knowledge in the form of evidence for and against the accuracy of a proposition. Truth values are replaced by the concept of a two-dimensional evidence space. We discuss the consequences of the concept, particularly with regard to verification. The connection between these two topics is established by implementation of the user model on the inference machine.

User Modeling
Broader fields of computer application going beyond routine data processing with stereotyped programs have, in turn, broadened the circle of computer users, and made it necessary to take different kinds of users into account in developing a system. Assuming a homogeneous user group, systems developers were able to design a system to perform in accordance with the requirements and capabilities assumed for a particular type of user (implicit user modeling). With a heterogeneous user group, this is no longer possible. The system is required to react differently to different users. In order to do this, it requires knowledge about one or more user types [a priori user models (Wahlster, 1982)] and rules which specify how this knowledge is to direct system behavior. Explicit user modeling is involved if, and only if, a system:

- Has access to explicit knowledge about users or user types, and
- Employs this knowledge in guiding system behavior.

In representing the a priori user models, earlier systems follow the approach of standards, i.e., representing the typical user explicitly (Brown and Burton, 1976; Genesereth, 1978) or the more flexible one of stereotypes, i.e., representing general knowledge about groupings of characteristics of persons (Rich, 1979).

In order to concretize an a priori user model for a particular user, i.e., to construct a user profile for a specific user from an a priori user models, two fundamental techniques have been developed. The system either takes the initiative and questions the user (Genesereth, 1978; Rich, 1979), or it directly determines a user profile from the user’s behavior by measuring deviation from a standard (Brown and Burton, 1976).

The user model we employ is based on the stereotype approach. It seems to us that this model has an important advantage over the standards approach because there are various a priori models for various aspects of individuals and a single a priori model is not required to uniformly cover all aspects of possible users. Thus characteristics that are not connected can be freely combined, and different views of an individual can be modeled.

In order to construct a concrete user profile, we utilize information supplied by the user about himself. Whereas the system in (Rich, 1979) asks questions about character traits, we only require information related to the topic
of the dialogue. Since self-assessments usually render a distorted picture of the user and are not expected in a real consultative dialogue, they should not be specially required in man-machine communication. In adapting a user model to an individual user, it is better to utilize only those characteristics that are mentioned by the users in the dialogue and which they consider to be relevant to their concerns.

At the level of the linguistic surface structure, certain phenomena like anaphora, choice of ellipses, and definite descriptions can be suitably handled with the help of user modeling (Wahlster and Jameson, 1982). On the conceptual level of communication, user models are necessary for the choice of suitable degrees of detail and the order of presentation of items of information (McKeown, 1982) as well as for the recognition of misconceptions (Webber, 1983).

Finally, on the action level, human characteristics have to be taken into account for the actual task of the system: here, to function as a consultant. The same solutions cannot be suggested to every user; rather, specific, individual solutions must be found. If a consultative system is required to do more than simply answer questions—and this is often in demand (Morik, 1983a)—it must be able to make suggestions and recommendations, to advise for or against a particular course of action. Since speech acts like advice, and recommendations are actions, the performance of speech acts is located on the action level as well.

User modeling is a necessary prerequisite for a suitable natural-language system response, but it is not only relevant for natural-language systems since it is independent of the concretely realized surface structure. The user model that we will present here operates on the action level. It is a core of a consultative system.

Value judgments can be seen as the central concept for the consultative situation. The speech acts of recommend, suggest, and advise are based on knowledge about the dialogue partner's value judgments. Something which he or she evaluated negatively cannot be sincerely recommended. Finding something suitable for the user requires knowledge about his or her assessments. The task of consulting may be seen as a comparison of the demands made on an object by the user and the system's knowledge about available objects, performed in order to determine the suitability of an object for the user.

It is often not possible to reach a clear decision about suitability. Let us suppose, for instance, that the system has the task of selecting a suitable apartment for the user from a list of offerings. It is not to be expected that any one apartment will meet all of the requirements of the user in every respect. In this case it is appropriate to consider the arguments for and against an apartment. After making a selection of the apartments that merit consideration, the system must make it possible for the user to assess those selections. In this case, the system cannot simply decide yes or no and make a positive or negative recommendation. Depending on the extent to which the advantages of a particular apartment balance out its disadvantages, the system must take different courses of action. It can make strong or weak recommendations, indicating which criteria are fulfilled and which are not, and leave the decision to the user. It may also note additional positive features of the apartment or call the user's unfulfilled requirements into question. This differentiated system response requires an internal representation of the user's value judgments (e.g., about apartments), i.e., the criteria that the user sets and the conditions for fulfilling a criterion. It demands a weighting of requirements, and it demands a comparison between the user's requirements and the characteristics of the objects (apartments), leading to a more differentiated response than a simple yes/no decision.

Our conception of assessment on a positive-negative scale is represented in Figure 1:

![Figure 1. Representation of Value Judgment Standards.](image.png)

Depending on the user profile, each criterion is assigned a specific importance value (very important, important, unimportant). For each criterion there are one or more conditions that must be fulfilled in order for the criterion itself to be fulfilled. Individuals differ not just in their choice of criteria but also in the conditions they impose for the fulfillment of a criterion (Morik, 1983b). For this reason, conditions assigned to a particular criterion are also dependent on the user profile.

Before we can describe how we implemented this concept of value judgment for user modeling on an inference machine which processes evidence, we must present the concept of evidence evaluation.

**Interpretation of Evidence-Evaluated Assertions**

Assertions about the world are by their very nature inex-
act. The inexactness can be broken down into incompleteness, uncertainty, and vagueness.

**Incompleteness:** As a first attempt at a definition we can propose the following: A world model is incomplete if the speaker/hearer does not have complete information about the extension of a relation or function (open vs. closed world). Since it follows from this that there cannot be any world model for a speaker/hearer which is complete, let us revise the definition so that completeness is not checked for the entire world model (the concept is meaningless at this level) but rather for a part of the world model. For example: “The apartment is in a quiet area.” The knowledge of the speaker is complete as far as traffic noise is concerned. It is, however, incomplete with respect to noise in general, e.g., the neighbors might be noisy. (See Collins et al., 1975, and Fox, 1981 on incompleteness).

**Uncertainty:** “The apartment will presumably be free on January 1, 1984.” The speaker uses the word “presumably” to articulate his degree of certainty about the likelihood that the event referred to will take place. He might be more or less certain that the event will take place than implied here. This would necessitate a different formulation corresponding to the degree of his certainty. The certainty or uncertainty of the event is the product of factors that speak for and factors that speak against the likelihood of its taking place. In particular, the world model relating to possible (future) events is systemically uncertain because the system of rules with which future world conditions can be derived is always subject to such uncertainty. (See Joshi, 1978, Lowrance and Garvey, 1982; Cohen and Grinberg, 1983; and Rollinger, 1983a, on uncertainty).

**Vagueness:** “The apartment has an area of about 100 square meters.” Here the speaker (the real estate agent) is quite sure that the statement is true. The element of inexactness arises through the characterization of the “object” (apartment). Moreover, it cannot be assumed unconditionally that the realtor’s world model is inexact on this point. Particularly in this context it must be assumed that the speaker would be in a position to make an exact statement (e.g., 105 square meters) but chooses not to do so, since, in accordance with Grice’s conversational postulates (“be relevant”), he does not wish to be overly informative. This criterion does not apply to uncertain knowledge since stating a certainty in a form different from the one actually available to the speaker in a situation in which the hearer cannot derive the actual certainty from context would be the equivalent of misleading the hearer and could not be explained with Grice’s postulates. The postulate “be honest” would be violated here. (See Zadeh, 1965 and Wahlster, 1981 on vagueness.)

In this article we will take up the uncertainty of statements. We believe that it is, in general, only possible to state the degree to which a proposition is likely to be considered true. This degree of certainty is determined by the quantity and quality of knowledge sources that speak for or against a statement. In order to determine the degree of certainty of a statement from the available information, the individual points of evidence must be combined, and the arguments weighed against one another.

A representation formalism for handling uncertainty must meet the following requirements:

- It must be capable of representing certain knowledge as well as uncertain knowledge.
- It must permit determining and explaining the reasons for uncertainty.
- It must permit comparison of degrees of certainty.

In the first two (Cohen and Grinberg, 1982), requirements are raised and a procedure completely dispensing with numerical values is presented. This neglects the third requirement, however. For this reason we utilize a mixed approach storing and applying both the degree of certainty in the form of numerical values and the reasons for the uncertainty.

![Figure 2](image_url)
for modeling human behavior, and we must assume that human behavior is not truth functional. In particular, this means that the different regions into which the evidence space can be divided cannot be interpreted truth functionally.

Let us turn our attention to examples (1) - (3):

1. AVAILABLE (APARTMENT -1, CARPETED) [1.0]
2. AVAILABLE (APARTMENT -2, CARPETED) [0.7, 0]
3. AVAILABLE (APARTMENT -2, QUIET) [0.2, 0.7]

The evidence values in these examples are to be interpreted as follows: [1,0] means that the person who has statement (1) in his knowledge base is convinced that this statement describes reality. (That is, there are very good reasons for and none against) In (2) with the values [0.7,0], the “owner” of the statement is “largely” convinced that the statement corresponds to reality, but is not completely convinced that reality completely corresponds to his model on this point. The uncertainty here may result from the fact that the source of information is the owner of the apartment, who was purposely vague about the floor covering in his description and did not explicitly mention the presence of “wall-to-wall carpeting,” although his description allows this conclusion to be drawn. In (3), there are arguments (points of evidence) that indicate that Apartment 2 is noisy (e.g., it is close to the street) as well as arguments that indicate that it is quiet (e.g., the presence of double windows). The figure does not indicate the degree of noisiness or quietness!

This makes it clear how the different regions of the two-dimensional evidence space are to be interpreted: The point [0,0] is understood to mean “no information about this matter is available”; the region [1,0] means “I am absolutely sure that this statement describes reality accurately (and completely)” The region surrounding [0.5, 0.5] indicates “There are arguments for as well as against,” and the region [0,1] indicates “I am absolutely convinced that the statement does not correspond to reality.” Finally, the region [1,1] means contradiction: “Every indication speaks both for and against the fact that this statement corresponds to reality.” As we will see below, other regions of the evidence space can also be meaningfully interpreted.

This sort of evaluation of statements and rules has certain consequences for the formulation of inference rules: truth values that can be used to define the concepts verification and falsification are no longer available in the evidence-evaluated statement space. These concepts must therefore be redefined. Corresponding to the points of evidence of the statements, we provide premises and conclusions with goal points with which the points of evidence of the statements, which are supposed to function as substantiation for premises, are compared. By means of this comparison, the deviation is the basis for the evaluation of the quality of the verification. The uncertainty of a rule is expressed by an indicator of its implicational strength a numerical value between 0 and 1.

A premise is verified only when it can be shown that there is a piece of supporting evidence whose evidence point lies in the area surrounding the goal point of the premise.

In order to calculate the evidence point of the conclusion, the deviations between evidence points and goal points of the premises are combined and multiplied by the strength of implication. The combinatory function can be expressed as the determination of a maximum, minimum, or an average. A goal point for the conclusion must also be established. The evidence point calculated for the conclusion is compared with its goal point. Hence a conclusion is verified only when the evidence point calculated for the conclusion lies in the area surrounding its goal point (usu-
Figure 4 illustrates an example of an inference rule. Involved here is one part of the construction of a user profile on the basis of two stereotypes: if a person is willing to pay a high rent, then the size of the apartment, the conveniences it offers, and its condition are very important criteria; if a person has small children, then whether there is a nursery school close by is, in addition, a very important criterion for renting an apartment. Here, the premises are supposed to be quite evident. The strength of implication is also set rather high with a value of \([0.9]\), and the conclusion should also reflect a high degree of certainty. The inference is triggered if the individual is prepared to pay a high rent, if the individual's children are small, or if it is to be shown that the above-mentioned criteria are very important for an individual.

### On the Modeling of the Action Level of a Consultative System

At this point we would like to describe the user model we implemented on the inference machine. We chose apartment rental as a typical consultative situation. The system assumes the role of the real estate agent, while the user is an apartment-seeker. Since we concentrated on the action level, the inputs and outputs were formulated either directly in the internal representation language SRL (Schneider et al., 1981) or in quasi-natural language with predetermined sentence patterns.

In inquiring about a suitable apartment, the user must first of all provide pertinent information about him- or herself—in this case, price range, the number of persons who will occupy the apartment, the presence of children and their age(s). These facts are assigned a high evidence value \([1.0]\) and entered into the system as “user facts.” Figure 4 gives an overview of the system’s knowledge bases.

The stereotypes, which are represented here in the form of inference rules (Figure 3), generate a list of which criteria are very important for the apartment-seeker, which are important and which are unimportant. The importance of the criteria is similarly entered under user facts.

For each criterion there are one or more alternative rules of application, which consist of conditions for the fulfillment of a criterion. Since users differ not only in the choice of criteria and their importance but also in the conditions set for the fulfillment of a criterion, there are different rules of application for a given criterion depending on user type. Thus an apartment may be classified as “large” on the basis of different rules of application:

- a) An apartment is large if it has as many rooms of a standard size as occupants.
- b) An apartment is large if it has more rooms of at

---

<table>
<thead>
<tr>
<th>CRITERIA</th>
<th>CONDITIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>STEREOTYPES</td>
<td></td>
</tr>
</tbody>
</table>

---

<table>
<thead>
<tr>
<th>USER FACTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>INPUTTED FACTS</td>
</tr>
<tr>
<td>IMPORTANT OF CRITERIA</td>
</tr>
<tr>
<td>APARTMENT FACTS</td>
</tr>
</tbody>
</table>

---

For purposes of implementation all figures are multiplied by 1000.
least standard size than occupants.

In this example the system chooses rule (a) if the user specifies a low price range and rule (b) if the user specifies a high price range.

A rule of application is formalized as a list of premises and the associated criterion as the conclusion. The conditions of a rule of application which are fulfilled by an apartment with a certain degree of certainty represent arguments for the suitability of the apartment. Conditions which are not fulfilled with an adequate degree of certainty represent arguments against the apartment. This is carried over to the criteria: if more arguments speak for the fulfillment of a criterion than against, this speaks for the apartment. In this way, an evidence value for the suitability of an apartment for a user is reached in two steps using the evidence computation.

The conditions correspond directly to possible characteristics of apartments Knowledge about apartments is represented in the form of SRL statements with evidence points. At present three different apartments are described by "apartment facts." The implementation of the concept presented above for value judgments along a positive-negative scale (Figure 1) can be illustrated using an example with two criteria, each with one rule of application (Figure 5).

In order to determine a suitable apartment for a particular individual, the metarule SUITABILITY is called. This rule attempts to verify the important and very important criteria for the user with the largest possible amount of evidence. If this procedure is not successful for any of the apartments, the important criteria are eliminated, and the very important criteria are checked again with a weaker grade of evidence. The result is then displayed depending on the evidence point determined for suitability (Figure 6).

The differentiation of possible cases, which is easily implemented using evidence points, provides a basis for modeling dialogue strategies and for speech act generation on the action level. If the suitability of an apartment for an individual has a high degree of positive evidence (a point close to [0,1]), then an unqualified recommendation of the apartment is in order.

If the criteria are fulfilled with a lesser degree of positive evidence (a point in the neighborhood of [0.6, 0]), a weaker (but still unqualified) recommendation is called for. A qualified recommendation is appropriate if there are explicitly unfulfilled criteria—that is, arguments which speak against an apartment but which are outweighed by advantages (a point close to [0.6, 0.3]).

If as much speaks for as against the suitability of an apartment (an evidence point in the region of [0.5, 0.5]), the user is in need of additional information. For this reason, the system searches for "extras" offered by the apartment. These are features which are not determined from the criteria set by the user, but which are generally assessed as positive. In this situation they may well tip the balance in favor of the apartment. An example here might be the presence of a fireplace (Figure 6).

Finally, if none of the available apartments fills many of the user's criteria, i.e., if negative evidence is very strong (a point in the neighborhood of [0,1]), there is no suitable apartment available to meet the user's requirements. At this point the system could be modified so that the criteria and conditions are outputted one by one and the user is permitted to modify them.

The user model takes into account the following assessment scales: “Quality” (good or bad), “importance” (very important, important, unimportant) and “evidence.” Evidence is processed with the help of the evidence space, supported by the inference machine. Importance is expressed by predicates. Quality is expressed in the form of rules involving value judgment standards (criteria and conditions). The relation between the scales can be pictured in the following way. The user's value judgment standards determine what is considered important and which features of an apartment are to be checked. The importance value determines the degree of certainty with which a criterion must be fulfilled. The degree of certainty with which the requirements of the user and the features of the apartment correspond determines the output behavior of the system.

Through differentiated treatment of different cases on the action level and through flexible construction of a concrete user profile based on several stereotypes, a basis is laid for a consultative system which utilizes the advantages of an inference machine employing evidence space. The concepts lend themselves to interfacing with processes on the conceptual level which are closer to natural language and with an NL surface structure. Thus, for example, the arguments for and against the suitability of an apartment could serve as a guide for the description of the apartment. The evidence points could control the choice of particles expressing conviction.

Implementation

The basis of implementation is an inference machine which was implemented in WPROLOG on the ITEL AS 5.3 Computer in the KIT Project as the nucleus of a text comprehension system. The fundamental characteristics of this inference machine are:

- The administration of an evidence-evaluated propositional knowledge base (here special reference should be made to the non-monotony of the inference process which can result from changes in knowledge, i.e., in connection with the use of default rules);
- The selection of inference rules both from the point of view of “rule complexity” and “compatibility of the evidence point of a proposition that is to be checked.
### Criteria Conditions Features of Apartments

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Conditions</th>
<th>Features of Apartments</th>
</tr>
</thead>
<tbody>
<tr>
<td>LARGE</td>
<td>LARGE AREA [8, 1]</td>
<td>AREA APARTMENT-1 LARGE [8, 0]</td>
</tr>
<tr>
<td>GOAL POINT</td>
<td></td>
<td>AREA APARTMENT-2 LARGE [0, 8]</td>
</tr>
<tr>
<td>LESS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>OCCUPANTS,</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IMPORTANT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>VERY</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IMPORTANT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CONVENIENCES</td>
<td>HAS WALL-TO-WALL</td>
<td>AVAILABLE APARTMENT-1 CARPETED [1, 0]</td>
</tr>
<tr>
<td>GOAL POINT</td>
<td></td>
<td>CARPET [1, 0]</td>
</tr>
<tr>
<td>IMPORTANT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>VERY</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IMPORTANT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ROOMS</td>
<td>ROOMS APARTMENT-1 FOUR [1, 0]</td>
<td>ROOMS APARTMENT-2 THREE [1, 0]</td>
</tr>
<tr>
<td>IMPORTANT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IMPORTANT</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Example for Two Criteria and Corresponding Apartment Facts.**

Figure 5.

...with the goal area of the entry point" in backward-chaining mode as well as forward-chaining mode;

- The capability of switching from backward-chaining to forward-chaining mode if a new (relevant) evidence value for a proposition is computed;
- A mixed depth-first breadth-first procedure oriented on the evaluations of applicable rules.

The next extension of the inference machine will take degree-of-interest assessments of conclusions into account as control knowledge in order to simulate inference processes directed by interest and attentiveness. The user model described here is a program consisting of 27 rules (of the type illustrated in Figure 3) with 43 entry points and 54 evidence-evaluated SRL statements.

This knowledge base (facts and rules) is interpreted by the inference machine. The system as a whole is started with a 1 megabyte storage. Answering a decision question or an additional information question formulated as an SRL expression requires, on the average, 2.7 seconds CPU time. This processing time is progressively reduced, since after being checked for consistency, all derived statements— including intermediate steps—are incorporated by the system into the knowledge base.

**Postscript**

After we implemented the kernel user model on the inference machine, we realized that the principles of user modeling described here are part of the natural language system HAM-ANS. In the hotel reservation situation, HAM-ANS builds up a user profile and infers the user’s evaluation standards. The system then accordingly recommends a room category to the user. Thus the embedding of this user model into a natural language system has in fact been implemented (Morik, 1984). However, revision of the assumed criteria and the user profile on the basis of the dialogue requires further investigation.

The inference machine has also been developed further. It has been enhanced following the principles of knowledge representation described by Brachman and Schmolze (1982). It now works entirely with an extended version of SRL; no PROLOG formulas are in evidence on the top level. Facts are no longer represented as a set of propositions but rather as a referential network with links to a conceptual network constructed by freezing inference rules (Emde et al., 1984). In addition, a clear interface for querying, updating and maintenance has been implemented.

**References**


tation von SRL J Laubsch (Ed ), GWAI-84 (forthcoming)
Input of user facts for user Maria

This entry corresponds to the content question “which apartment is suitable for Maria?” The symbol “*” signifies a variable. If an apartment can be found the statement is stored under *fact

A weak but unqualified recommendation for apartment_2 is given. This natural language answer is generated with sentence patterns.

This is the internal result of DEDUCE. *apartment is instantiated with apartment_2, the evidence evaluation of the proposition is stored under f67, the instantiation of *fact, together with the rule with which it was possible to derive f67 and the facts used in the derivation.

Let’s see f67

The evidence point 729 0 for the proposition (suitability(apartment_2,maria)) was determined with the aid of facts 59,63,63,64, and 65 by rule 3.

Input of user facts for Paul, who has a small child.

The natural language equivalent of this question is “which apartment is suitable for Paul?”, once again a content question.

Clearly, apartment_1 is better suited for Paul than apartment_1 is for Maria.

A decision question “Is apartment_3 suitable for Paul?”

The low level of suitability of apartment_3 for Paul occasions the system to seek out extras (a fireplace) which make the apartment more attractive. The extra is outputed.

The system could not find an evidence point for the proposition which is compatible with the implicitly demanded point [1,0]. Hence no facts were entered into the knowledge base. This produces the output of a question mark. That the system generates an answer despite this is due to the fact that DEDUCE triggers a backward chaining inference, which also determines an evidence point for the proposition (at least [0,0]), and that—depending on this evidence point—a forward-chaining inference is triggered which seeks out extras and, if it finds one, generates the answer.

End of dialog.

Figure 6. Translated and commented but unedited example session.