Dialogue Managment with weak speech recognition : a pragmatic approach

Florian Seydoux, Alex Trutnev, Martin Rajman

Artificial Intelligence Laboratory,
Institute of Core Computing Science,
Swiss Federal Institute of Technology, Lausanne
http://liawww.epfl.ch

Abstract

The present contribution adresses the design of pragmatic solutions for various problems occuring within dialogue-based vocal systems using low performance speech recognition engines (SREs), for example in situations where the used speech recognition engine is not fully adapted to the specific application, or the data necessary for training reliable acoustic and language models is not available. To accomodate the use of a low performance SRE, the following design principles are used to guide the conception of the dialogue model: (1) adopt a (limited) mixed initiative dialogue management strategy to improve flexibility of use; (2) avoid repetitions in the dialogue flow; (3) integrate in the dialogue management strategy mecanisms for recovering from specific dialogue repair situations such as request for help, for repetition, miscommunication, ...; (4) minimize the duration of the dialogues, i.e. aim at dialogues providing the user with the relevant information in a minimal number of turns; (5) provide the user with adequate feedback information about the state of the dialogue and the recognized pieces of information; and (6) filter out as much conflicting data as possible.

The structure of the contribution is the following: we first detail the context in which our dialogue model was designed; then we describe the solutions that have been proposed to implement the above mentioned design principles. Next, from the final evaluation of the system, we derive some insights on the impact of the selected solutions on the user perception of the system. All the proposed solutions were designed, implemented and evaluated in the framework of the *InfoVox* project.

1. Context of the project

The *InfoVox* project[1] was jointly realized by EPFL, IDIAP, and the Swisscom and Omedia companies. The main goal of this project was the elaboration of a methodology for the rapid prototyping of vocal information servers [1, 2], and the application of such a methodology to develop a prototype for vocal (and Web) access to information about the restaurants in the city of Martigny, Switzerland (e.g. [3]).

One of the main problems occuring during the implementation of information servers is the management of the interaction with the user. Indeed, for simple and well structured tasks (e.g. phone box management, credit card information services, book rental, etc.) or for applications designed for "trained" users, standard technics based on menu driven dialogues relying on option selection with DTMF keys, or a restricted set of keywords, appear as quite sufficiant. However, such very simple technics are clearly not adequate for more complex applications, as for example, the ones to search an element by some criteria with many modalities (travel booking agent, search for books, etc). It is necessary for these applications to increase the ability of understanding, in order to make the interaction with the users more natural.

Implementation of such "interaction" implies capacities much higher in terms of speech recognition (and possibly, text-to-speech), as well as the real behaviour of "dialogue". From this point of view, the main problem (at least, for tasks of relatively low complexity) is the detection and management of miscommunications that could happen between the user and the system.

These miscommunications, although unavoidable for natural language interactions, are however strongly influenced (in term of frequency, detection and management) by the quality of the speech recognition system.

In our case, the used speech recognition engine appeared to be of a very poor quality. Implementation of the technics assuming the robustness of dialogue became necessary. The idea here was on one hand to detect as many as possible of cases of miscommunication, and on the other hand to limit the number of those cases.

1.1. Prototype overview

As Web access to the application is an easy task, the main efforts concerning the prototype were essentially concentrated on vocal access. The objectives to be achieved in the prototype were :

dialogue in French

natural intuitive input even an unexperienced user should be able to use the system, pronouncing his/her needs in a natural way; an experienced user should be able to express his/her needs in direct and natural way. This implies that the system operates on unconstrained language, and that information gathering aiming to realize the task must be done in different ways.

(limited) mixed-initiative interaction dynamic exchange of the control flow has to be defined in such a way that either user or system can easily guide the dialogue; contrary to many comparable systems (e.g.[3, 4]), such a mixed-initiative has to be defined not only for corrections or clarifications, but also for the whole way of realizing the task.

robustness and effective grounding the dialogue manager should also operate in spite of speech recognition errors (this point is particuliary important in *InfoVox* – c.f. 1.2) and "be sure" about the mutual understanding of both the user and the system during the dialogue. Namely,

[1] Partly funded by the Swiss national CTI grant program

> the problems of miscommunications have to be managed (treated) as much as possible in the framework of the dialogue, and this always in the most "natural" manner.

The general architecture of the prototype is resumed on the figure 1.1. The components interesting for this contribution are briefly described in the following sections.

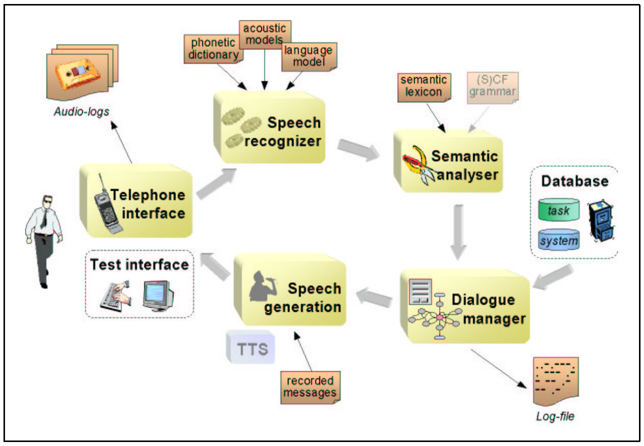

Figure 1: *System architecture*

1.2. Speech recognition

The system was designed for multi-speaker continuous speech recognition and its final evaluation showed a weak performance of its speech recognition component. Two main reasons explaining this are connected to acoustic and language models. As to acoustic model, the environnement in which operated the system was not the same as the one used to train acoustic model (more noisy telephone lines, spontaneous speech during field-tests vs written speech used for training). Moreover, evaluation of the acoustic model itself showed its weak performances to estimate. Language model was trained on the data acquired during the Wizard-of-Oz experiment of the project, but the data (15’000 words covering 1’000 lexicon words) was not sufficiant to estimate reliable language model. Unfortunately, neither time nor manpower in the project was sufficient to train more models. The dialogue model was thus conceived (and modified) in a way that it deals with these weaknesses.

The developed SRE adopted the hybrid approach, in which estimation of probabilities distributions of phonemes is made with a neural network, and decoding is performed with hidden markov models ([5]). However, despite the fact that the underlying methodology is considered as robust and performant, the evaluation of the SRE showed its poor quality. In the table 1 are presented the main results of this evaluation. Several state-of-the-art SREs were used as references ([6]). The data used for evaluations were acquired during the internal and external field-tests of the system (3).

1.3. Speech generation

The module of speech generation developed in the framework of the project is reduced to a set of predefined messages, some of them complete, others in form of segments combined by the dialogue manager.

System	Internal Field-test, **WER**, %	External Field-test, **WER**, %
InfoVox	85.6	87.4
HTK	61.5	63.3
Nuance	65.5	65.0
Loquendo	67.9	66.3
Sirroco	66.6	68.9
ViaVoice	71.0	75.2
Noway	72.5	76.6

Table 1: Evaluation SREs on *InfoVox* data

1.4. Semantic analyzer

The role of semantic analyse is to find, in the utterance returned by the speech recognizer, key or key sequence defined for the application, and to replace them with a triplet of the form <*context, semantic value, effective value*>.

1.5. Dialogue manager

The architecture of the dialogue manager is composed of a mix of *finite state script*, *frame-based* and *sets of contexts* ([7]).
Namely, the dialogue manager is a finite state machine composed of action nodes and generic dialogue nodes (GDN), whose role is to fill several questionnaires. Every questionnaire corresponds to a part of the task that can not be realized in parallel with the others (for example, one restaurant has to be proposed to the user, before proceeding to the next restaurant).
More precisely, the task is modeled as a set of frame in which the fields with associated contexts represent the various attributes that need to be informed for the task to be performed. For example, for our application, the following fields were used to model the search subtask: ’type of food’, ’slice of price’, ’localization of the restaurant’, ’open days’, ’timetable’, and other subtasks are essentially associated to "yes/no" questions (corresponding to a trivial questionnaire), and permit to provide user with the asked information, in the second part of the dialogue.
One advantage of this approach is that the interface for Web access can be derived directly from the dialogue model; in addition to the frames, an input line can be used to enter free text[2], directly connected to the input of the *semantic analyzer*[3].

2. Implementing the dialogue manager

In order to achieve the objectives stated above, a set of principles, obtained from analysis of the results of Wizard-of-Oz [1], was used to guide the developement:

1. (limited) mixed-initiative interaction;
2. avoid repetitions and heavinesses in the dialogue;
3. deal with the *dialogue repairs*[4] during the dialogue;
4. minimize the duration of the dialogues;
5. feedback about the state of the system;

[2] Of course, a posting zone and the possibility to switch to different frames should also be defined.

[3] In fact, vocal access works in parallel with Web access, sharing the same dialogue manager.

[4] Term used as reference to *speech repairs*, describing here dialogue acts related breakdown and repairs sequences on the dialogue (*effective grounding* in [8]).

6. dealing with misconceptions (conflictual informations).[5]

Next sections present how these principles have been implemented during the realization of the prototype .

2.1. (limited) mixed-initiative interaction

The implemented system permits in fact only a *limited mixed-initiative interaction* ([4]), consisting in the possibility for the user on one hand to break the dialogue flow imposed by the system by asking for repetition or explanation of the last question (cases assimilated to *dialogue repairs*), and on the other hand to anticipate the future questions by providing the elements of response in advance [6].

Note however that the user can indeed choose not to answer the asked question: on one hand, it is possible that the additional informations provided by anticipation result in realization of the current task, and thus progressing in the graph of the state machine; on the other hand, as the attribution of a field to a dialogue node is realized dynamically, during every visit of the node, a loop on this node will not necessarily lead the system to ask the same question.

In consequence, as in the case of a dialogue between humans, an unexperienced user will be guided by the system, while an experienced user will be able to provide the system with the pertinent elements at once ([10] , but it's not necessary, in our case, to explicitely define the shortcuts).

2.2. Avoid the repetitions

In order to avoid too "mecanical" aspect of the system during the dialogues, it appeared important to avoid the repetition of the same system messages.

For this purpose, different alternative formulations, more or less equivalent, have been defined for each system message; in addition, some of them were contextualized (for example, welcome and good-bye message, according to the hour and the day of the dialogue : "good evening", "have a nice day/week-end").

In the cases where one message has to be played several times during the dialogue, a formulation that differs from the most recent occurences is choosed.

Beside more natural aspect, this approach can be implemented as mecanism of disambiguation, when the user indicates that he doesn't understand the message.

2.3. Dealing with the dialogue repairs

For an answer of the user, the following situations should be considered, on the basis of values and contexts found in the answer, as well as in the context imposed by the system:

Ok + Initiative *The user answers the question, or the user answers something, which the system can interpret (even partially) as an answer to an different question, and can take into account as an ok case.* This case is not a *dialogue repair*, and the dialogue can continue.

Repetition *The user asks explicitly for the repetition of the system last utterance.* The last system message will be repeated. When the repetition is asked for several times (2, 3), it is preferable that the system gives an alternative formulation.

User misunderstanding *The user says (more or less clearly) that he did not understand the question.* An alternative formulation of the last system message can be used at the first time, and in the case of new misunderstanding, the system continues with the *request for assistance*. When the concerned message is an opened question, and the misunderstanding is repeated, the system continues with the guided dialogue, in which the system asks the user to inform only one field of the frame.

Request for assistance *The user does not know how to answer the question, but keeps in mind that he is faced with a machine.* The system indicates to the user how to answer, by providing him with valid example of answer, or by closing (at least partially) the question.

Timeout *The user stays mute after few second (recording stop on time-out).* The system asks the user to talk louder and to wait for the signal before answering, then continues with the *repetition*.

System non-understanding + Out of context *Nothing can be extracted from the answer (problem upstream the data processing sequence (recognition), overcomplex answer, rumble and other non verbal, etc.), or the user answers something that has no relation with the task; this scenario not being treated by the system,[7] it will be automatically assimilated to a case non-understanding.* The system indicate to the user that he was not understood, and require a repetition. If this situation occurs again, the *request for assistance* is triggered; as for the *user non-understanding*, the system can then switch to driven dialogue, or use some exit mechanisme (based on watchdog) to go out of the loop.

2.4. Minimization of the duration of dialogues

The minimization of the duration of the dialogue (in terms of dialogue turns and total time), without any restriction on the task, is in fact more an objective than a way to improve the dialogue[8].

Many factors can be cited that influance the duration of the dialogue. However, following considerations are always useful:

- For long messages that can be played several times during the dialogue, it can be useful to dispose, in addition to alternative formulations, shorter (elliptic) reformulations. Multiple long formulations can be used to clarify the main message and its reformulations in the case where the same message has to be played later during the dialogue.
- Mixed initiative is also an efficient way, because it permits to indicate several informations at a time.
- Generally, loops have to be avoided in the dialogue. This is not necessarily easy to realize, because some loops are mandatory, for example, when the user asks explicitly several times the system to repeat.

  The solution retained in the framework of *InfoVox* was to implement the control at several levels:

[5] For a classification of *misconception*, *misunderstanding* and *non-understanding*, see [9].

[6] This is typically the case with the half-opened question serving to initiate the dialogue.

[7] Indeed, this treatement requires the modelization of the "contexts" outside the application (word knowledge); this could be however foreseen for a well defined set of contexts, defined on the basis of *a posteriori* examination after sufficiently long period of functioning of the system, and for which this treatement would be justified.

[8] For the same result, more short dialogue is more "efficient"; one limits the risk to annoy the user, as well as possible miscommunications.

- in the case of consecutive non-understandings (user or system) of an opened question, a driven dialogue is instanciated;
- for the same situation, but in a driven dialogue, for a frame containing several empty fields, the system gives up with the current field and switches to the next empty one;
- in the cases of loops on non-understandings, but with the questions "yes/no", the system selects a (default) value inducing as less consequences as possible, and plays a message to the user explaining such a decision;
- global watchdog, instanciated in the case of non-progressing in the dialogue (for example, when no field is informed or modified during last n interactions, strict repetitions excluded); in this case, depending on the application, different actions can be considered: either one proposes to the user to present already available informations, or user is proposed to abort the dialogue, possibly contacting the human operator, etc.[9]

In addition, one can also profit of the nature of the task, either by choosing "intelligently" the questions to ask (prior fields to be informed first), or, in case of troubles, terminating the task by providing some partial information (or service):

search task *information retrieval of one or several elements a priori present in the database.*

For this kind of task, one can choose, for a driven dialogue, to inform in priority the fields with the highest potential to discriminate; this will minimize the average duration of each dialogue (an algorithm based on ID3 is proposed on appendix B; see also [11]).

An additional advantage of this technique is that it offers the possibility to weight the elements of the base depending of their "popularity", permiting to minimize even more the average duration of dialogues[10].

advice task *search for an advice or propositions; it is still information retrieval, but of an element that is not necessarily in the database*

This strategy has been retained in *InfoVox* (the system proposes a set of restaurants on the basis of preferances indicated by the user). In this case, the discriminating potential for a question has no sense any more. On the other hand, it appears possible to propose to the user a set of elements that do not satisfy the whole set of attributes, but only a sub-set of them.

This technique is relatively easy to implement: the user indicates her/his choices during the dialogue. When no database element satisfies the indicated choices, rather than to terminate or restart the dialogue or propose to the user to modify some choices, the system can ask if the user agrees to accept a sub-set of her/his preferencies. It is possible then to relax the last specified constraints,[11] possibly in successive way (in this case be careful not to submit several times the same propositions).

Before proposing to the user the "approximate solutions", one has to be sure that the constraints relaxation provides the solutions that conserve a sens in regard to the user request; if this technique is applied in the cases where the user indicated just few constraints, the relaxation of one can lead to huge number of uninteresting elements. In the framework of implemented prototype, one controls, before proposing the solutions, that the number of targets obtained after each relaxation remains reasonable compared to the number of excluded targets. In the cases where this ratio becomes too big, the mechanismes cited above (ask the user to modify the request, submit it or terminate) are applied.

2.5. Feedback about the state of the system

In the case of dialogue between humans, the progression in the dialogue is accompanied by acknowledes. On the basis of a corpus of spontaneous conversational speech of about 200,000 interactions, [12] found that about 20% of dialogue acts were acknowledges.

These acknowledges were not implementable in the framework of *InfoVox* (no barge-in, too hazardous recognition of acknowledges, reaction time of the system, etc.).

In order to avoid that *system misunderstandings* (numerous because of speech recognition errors) don't lead the system to provide the user with solutions without any correlation with her/his needs, it is necessary to detect and to handle these situations; the only possibility to do it is to inform the user about the retained elements.

The first implementation of such feedback is, before terminating a sub-task (for example, at the exit of a frame), to produce a synthesis of different retained elements (the fields values recently informed after the last synthesis).

However, this technique alone was considered not sufficient (the correction of erroneous values can lead to tiresome dialogues, the users are not always attentive enough to realize that the system misunderstood what they said, especially when the number of fields is important, ...).

It was therefore decided to implement a kind of acknowledge during the dialogue, by including, as perfix of each closed question, an indication of one of the fields[12] recently informed.

Here is an example of system message composed of confirmation prefix and question about the next field: "For your meal tomorrow night, what localization do you prefer?"

In the case of agreement (or no explicit protest) from the user side, the system marks the field as being "confirmed"; it will not be possible from this moment to modify it by mixed-initiative mecanism. If the user protests, either by negative response, or by indicating a value compatible with (but different from) the

[9] The dialogue can become untreatable in the case where non-understadings arrive at this moment. A good prevention is then to use a minimaliste dialogue, by using DTMF keys.

[10] However, one should to control that the scenarii too surprising and able to destabilize the user are not choosen, for example, ask the user about the color of an article, before asking for the nature of this article In addition, mixed initiative constraints to realize this choice dynamically; if the database is huge, the number of attributes is big and the attributes are highly various, this choice can become relatively expensive operation.

[11] Counting that first indicated elements are most important.

[12] Several possibilities to choose a field can be considered: the first informed, the last informed, etc. In the framework of *InfoVox* project, the retained criterion is to choose the fields of the most recent answer, depending on the order of appearing of the values.
One has in addition to consider the context of the question while choosing the fields for confirmation, avoiding the couples with shared modlities.

confirmed one, the system asks the user for her/his preference about the field, by prefixing the possibly conflict values.

This solution can not guarantee the confirmation of all informations that will be used by the system, as is the case in [13], especially if the user anticipates the questions of the system, but it still remains balanced between a limited negative impact on the dialogue quality and the most prompt correction of recognition errors.

2.6. Dealing with misconceptions (conflictual informations)

When the user indicates several incompatible values for the same field, a sub-dialogue aiming to define the desired value is instanciated (the system asks the user to choose one value for that field, and indicates in addition the conflict elements).

To avoid that the recognition errors do not lead to these conflict situations, the informations identified by the system are strongly filtered, before the fields are update:

- If the response of the user is given in the context imposed by the last system message or contained in the confirmation feedback prefix, only the elements compatible with one or other of the contexts are conserved. In the case where no such elements exist, the filtering is canceled and all the elements of the response are taken into account.[13]
- If the analysis of the user response indicates the presence of at least one non conflict value, the conflict elements are filtered, except for those the context of which is compatible with the possible confirmation feedback.

3. Evaluation of the system

The evaluation of the prototype was essentially carried out on the basis of two (internal and external) field tests[14]. For the external field test, a population of 50 "external" users (i.e. users that did not have any a priori knowledge of the system) was randomly selected in all French speaking cantons in Switzerland.

Several subjective and objective indicators have been derived from the raw data produced during the test. Subjective indicators essentially corresponded to average scores obtained for the various closed questions present in the satisfaction questionnaire, while objective indicators have been derived from the logfiles and corresponded to average measures of various system characteristics (such as Word Accuracy, Word Error Rate, interaction duration, number of Help requests, ...) describing its interaction with each of the users.

Concerning the exploitation of the produced indicators, 3 kinds of analyses have been carried out:

Retrospective trend analysis *identification of the subjective indicators corresponding to significantly predominant modalities of some closed question, and can be used to provide retrospectively a synthetic view of the opinion of the users about the system.* Two main conclusions regarding the whole prototype was that the average global satisfaction was of 63.75, the system was seen as easy to use (89.8%). Concerning the problems adressed above, the interaction duration was adequate (72.9%), the sequencing of the questions was considered as natural (93.9%), the users rarely (14.0%) felt lost, the majority (79.2%) of the users were sensitive to the confirmation messages and considered (96.8%) such confirmations as useful. Finally, concerning dialogue initiative, no clear opinion emerged with respect to predominance of system- or user-initiative.

Retrospective correlation analysis *identification of significant correlations between the answers to pairs of closed questions.* Two main conclusions here were that users having considered that the system was not producing correct results showed a significant tendency to consider the system as non satisfactory, and that there was no significant correlation observed between user satisfaction and the subjective indicators related with the readiness of the users to use or recommend the system.

Prospective correlation analysis *identification of significant correlation between the answers to some closed question and some objective indicator derived from the logfiles in order to prospectively guide the identification of promising modifications of the existing prototype that could lead to better user satisfaction.* The most important consluion made during this analysis was that in order to increase user satisfaction, it is important to act in priority on the quality of the interaction (at the expense, for instance, of the improvement of the background module producing the system results).

4. Discussion - conclusion

The work has shown that the proposed methodology is able to deal with speech recognition errors without specific need for huge linguistic resources. More reliable speech recognition is of course more preferable, but the solutions implemented in the framework of *InfoVox* are still interesting, because one can say that the speech recognition errors are unavoidable for any dialogue system based on vocal interactions. Important note concerning this work is that the proposed methodology of conception and developement of dialogue systems is “generic”, contrary to different projects realized in the past. The solutions aiming to deal with speech recognition errors presented in this contribution (namely, limited mixed initiative, dealing with the repetitions, dialogues repairs and conflictual informations, confirmation strategy and strategy for minimization of the dialogue duration), are integrated in a “natural” way in the methodology. The whole methodology was implemented for a simple task with restricted and discriminative dictionnary; next step is to deploy such architecture and proposed solutions for more complexe applications, as for example, *INSPIRE* (aiming at a dialogue based control of various home devices (lights, TV, VCR, ...) within a Smart Home environment), or *IM2*, aiming at the set up of efficient dialogue based interaction mechanisms with a database of multimodal meeting transcriptions.

[13] Such filtering is relatively rough and strongly limits the initiative of the users. Better solution would be to filter only the elements for which the confidence score is too weak. This would also permit to short-cut the filtering for the Web interactions.

[14] A detailed description of the evaluation and its major results can be found in [1]

A. References

[1] M. Rajman, *et al.*, "Assessing the usability of a dialogue management system designed in the framework of a rapid prototyping methodology," in *Proceedings of the 1st ISCA Workshop on Auditory Quality of Systems*. Mont Ceni, Germany: ISCA, april 2003.

[2] R. V. Kommer, M. Rajman, and H. Bourlard, "Heading towards virtual-commerce portals," *Comtec*, pp. 10–13, sep 2000.

[3] D. Albesano, *et al.*, "Dialogos: A robust system for human-machine spoken dialogue on the telephone," in *Proc. ICASSP '97*, Munich, Germany, apr 1997, pp. 1147–1150.

[4] J. Allen, G. Ferguson, and A. Stent, "An architecture for more realistic conversational systems," P. of Intelligent User Interfaces 2001(IUI-01), Ed., Santa Fe, jan 2001.

[5] H. Bourlard and N. Morgan, "Continuous speech recognition: An introduction to the hybrid hmm/connectionist approach," *IEEE Signal Processing Magazine*, vol. 12, no. 3, pp. 25–42, May 1995.

[6] A. Trutnev, "Evaluation of speech recognition engines," EPFL, Switzerland, Tech. Rep., to appear in 2003.

[7] J. Allen, *et al.*, "Towards conversational human-computer interaction," *AI Magazine*, 2001.

[8] J. Allen, *et al.*, "An architecture for a generic dialogue shell," *Journal of Natural Language Engineering, special issue on Best Practices in Spoken Language Dialogue Systems Engineering*, vol. 6, no. 3, pp. 1–16, dec 2000.

[9] G. Hirst, *et al.*, "Repairing conversational misunderstandings and non-understandings," *Speech Communication*, vol. 15, pp. 213–230, 1994.

[10] L. B. Larsen, "A strategy for mixed-initiative dialogue control," in *in proceedings EUROSPEECH -1997*. Rhodes: EUROSPEECH, 1997, pp. 1331–1334.

[11] J. R. Quinlan, "Induction of decision trees," in *Machine Learning*, 1986, vol. 1, pp. 81–106.

[12] A. Stolcke, *et al.*, "Dialogue act modeling for automatic tagging and recognition of conversational speech," *Computational Linguistics*, vol. 26, no. 3, pp. 339–373, 2000.

[13] M. Danieli, "On the use of expectations for detecting and repairing human-machine miscommunication," in *Proceedings of AAAI-96 Workshop on Detecting, Preventing, and Repairing Human-Machine Miscommunications*, Portland, OR, 1997, pp. 87–93.

B. Most discriminating field

Different attributes characterizing the targets and used for the selection serve as criteria permitting the partitionning of the search space into different classes.

Depending on the repartition of the values of these attributes on the set of not yet isolated targets, some fields offer more advantages to be informed than others. Thus, in Switzerland, where almost all Chinese restaurants are expensive, the knowledge about the price is of less interest than any other field, for example, localization, if the user indicated that she/he was looking for a Chinese restaurant. Consequently, as far as the element one looks for is in the database, a correct choice of order of information of the field can lead more quickly to a solution or a set of potential solutions (this set being small enough to be given to the user.)

From the point of vue of information theory, one will say that the informative value of knowledge is different for each field. In order to minimize the number of interactions with the user, it would be preferable thus to choose first the fields with the bigest informative values, i.e. the fields resulting in partitionning the targets that minimize the disorder (entropy).

The *informative gain* for the set of potential targets Ω (i.e. not yet isolated targets) brought by the knowledge of the attribute γ is measure by the reduction of uncertainty $I(\Omega\,;\gamma)$:

$$I(\Omega\,;\gamma) = H(\Omega) - H(\Omega \mid \gamma) \tag{1}$$

To maximize the informative gain, one will select to inform the attribute γ that minimizes entropy $H(\Omega \mid \gamma)$ [15]

with

$$H(\Omega\,|\,\gamma) = \sum_{m\in\gamma} P_\gamma(m) \cdot \underbrace{H(\Omega\,|\,\gamma{=}m)} \tag{2}$$

$$H(\Omega\,|\,\gamma{=}m) = -\sum_{\omega\in\Omega} P_{\Omega|\gamma}(\omega\,|\,m) \cdot \log\left(P_{\Omega|\gamma}(\omega\,|\,m)\right) \tag{3}$$

The following algorithm – specially adapted for data handled by external database – can be used to determine, on the basis of the simple criteria, the field to be informed in priority.

Algorithm B.1

```
***BestChoice***$(\Omega, \Gamma)$ :
  $\Omega$ = *targets selected by the fields already set;*
  $\Gamma$ = *remainding fields;*
*Step 1: initialisation*
  **foreach** $\gamma \in \Gamma$
    **foreach** $m \in \gamma$
      $\Upsilon[\gamma][m] \leftarrow 0$
*Step 2: scan database*
  **foreach** target $\omega \in \Omega$
    **foreach** field $\gamma \in \Gamma$
      **foreach** modality $m$ available into $\omega \Rightarrow \gamma$
        $\Upsilon[\gamma][m]$ += 1
*Step 3: compute the criteria*
  $H_{min} \leftarrow \infty$
  **foreach** field $\gamma \in \Gamma$
    $H_\gamma \leftarrow 0$
    **foreach** modality $m$ allowable for $\gamma$
      $H_\gamma$ += $\log_2(\Upsilon[\gamma][m]) \cdot \Upsilon[\gamma][m] \cdot |\Omega|^{-1}$
    **if** $(H_\gamma < H_{min})$
      $H_{min} \leftarrow H_\gamma$
      $\psi \leftarrow \gamma$
***BestChoice*** $\leftarrow \psi$
```

[15] To avoid that the attributes with many modalities are systematically choosen (for example, an identifiers have a nul uncertainty, but are not very interesting), one will prefere to minimize $\frac{H(\Omega|\gamma)}{|\gamma|}$