SCHISMA: A NATURAL LANGUAGE ACCESSIBLE THEATRE INFORMATION AND BOOKING SYSTEM.

G.F. v.d. Hoeven, J.A. Andernach, S.P. v.d. Burgt, G-J.M. Kruijff, A. Nijholt,
J. Schaake, and F.M.G. de Jong.
Parlevink Group
University of Twente, Dept. INF/SETI
P.O. Box 217
7500AE Enschede, The Netherlands.
e-mail: {vdhoeven,andernac,stan,kruijff,anijholt,schaake,fdejong}@cs.utwente.nl

Abstract

This paper gives an overview of activities in the SCHISMA project now and in the near and further future. Main points of discussion on current activities are collecting data on dialogues, dialogue analysis and dialogue modelling. Some attention is also paid to experiences using a the commercially available product Natural Language.
In the discussion of future plans and developments, topics such as speech, integration of speech and language, and the possible contribution of neural networks and fuzzy logic in the project are considered.

1. INTRODUCTION.

SCHISMA is a collaborative project of the University of Twente and PTT Research. The participating groups are the Parlevink group of the Department of Computer Science in Twente, and the speech and language group at PTT Research.

The aim of the project is to develop a prototype of a natural language dialogue system. The envisaged system is capable of providing a user with information about theatre performances, and it should allow the user to book seats for such performances. In addition to the goal of building a prototype of some quality there is the equally important goal of gaining a deeper insight in the problems one encounters in the process of building a natural language dialogue system. Getting experienced is considered one of the prerequisites for a successful follow-up of the project.

In SCHISMA *dialogue management* has been identified as the key issue to be addressed. As a consequence emphasis is put on the proper combination of the various sources of information, on the current status in a dialogue, on the process of judging that status, and on the subsequent steps. Parsers, semantic representation formalisms, knowledge representation formalisms, etc. are tools to be used, rather than products to be delivered. Of course improving these tools could also lead to improvement of the capability of our system. But in our view the key to enhanced functionality and better performance lies with better dialogue control.

Characteristic for SCHISMA in comparison to projects with similar objectives is the choice of the application domain: theatre performances and bookings. Furthermore one could probably characterize SCHISMA as 'eclectic': it is tried to find the proper mix of ideas, new ideas are developed only when necessary, and special attention is given to integration.

In this paper an overview will be given of major issues in current SCHISMA research. The project plan distinguishes between three related lines of research and development, each with their own deliverables.

1. **The collection and analysis of a corpus of real dialogues.** This aspect of the project will be addressed in sections 2 and 3 which contain a discussion of the Wizard of Oz environment that was built, a discussion of various forms of dialogue tagging, and of our views on obtaining a useful corpus of annotated dialogues.
2. **Investigating notions of dialogue state and state transitions.** Next topic, discussed in sections 4 and 5, is dialogue control and functionality of a natural language dialogue system. Under this heading there is attention for on a suitable notion of discourse semantics and representation,

and for the possible role of discourse transition matrices.

3. **The realisation of a theatre information system with the use of a commercially available tool set for building natural language interfaces to databases.** As is discussed in section 6, the exploration of the possibilities and shortcomings of available tools was considered useful. In addition, it could provide a possibility for a comparison with the envisaged SCHISMA prototype and demonstrate the added value of real dialogue over question answering.

The paper concludes with an outline of future developments in section 7. Here it will be argued that SCHISMA is suited for aims which are less restricted than the current aims formulated above, and that it may turn into a much broader research umbrella, with more and continuous attention for issues like speech, speech and language integration, robustness and neural and fuzzy approaches.

2. The Wizard of Oz experiments for SCHISMA.

As mentioned above, one of the SCHISMA objectives is to conduct Wizard of Oz experiments in order to obtain data on dialogues between a human user and a (simulated) automated natural language dialogue system. Stated more precisely, the objectives of the experiments are threefold:

1. to gather dialogues which can serve as a basis for the design of a dialogue manager for our prototype man-machine dialogue system
2. to develop and refine dialogue models derived from dialogues obtained from a relatively free subject-wizard dialogue
3. to find out how the subject-wizard dialogues are evaluated by the subjects compared to normal man-man dialogues, with regard to robustness and user-friendliness.

2.1. General description.

In order to conduct the experiments, one needs people to participate and a wizard environment to work with. The involvement of humans will be discussed first.

There are four actors involved in each experiment session: a subject, the subject's instructor, the wizard's assistant and the wizard. The instructor and the assistant have auxiliary roles. The main roles are played by the subject and the wizard.

The subject is the person who acts as (or is) a client using the system. The subject is not aware of the fact that there is no real automated system, and that (s)he is communicating through the Subject's Interface with a person that provides the system's functionality. The number and kind of subjects depend on the experimental phase (see below).

The wizard is the person who simulates the functionality of the system. The simulation involves

- the elicitation, recognition and understanding of the user's utterances,
- acquiring the data needed to query the database,
- interpreting the output of the database query,
- formulating adequate responses to the user's utterances.

The wizard is one of the SCHISMA researchers, who is supported in his task by the Wizard Interface.

Conducting dialogues in a wizard environment through a wizard has an additional goal, which involves especially the wizard and his assistant. Besides simulating the system's functionality, the wizard and his assistant are also supposed to monitor the wizard's behaviour in order to provide input for the design and development of the automated dialogue manager. One way to obtain relevant data in this respect, is to let the wizard work under various restraining conditions. For example, the wizard can be given the instruction to follow a restrictive protocol that forbids him to reply to a question with another question, or to select all his reactions from a very restricted set of possibilities, etc.

Another parameter in the experimental setting, is the wizard environment. It consists of the Wizard Interface and the Subject Interface. The Wizard Interface helps the wizard to control the dialogue simulation session. The Subject Interface allows the subject to have a keyboard conversation with the wizard.

More specifically, the Simulation Environment allows both the wizard and the subject to enter, edit and transfer their utterances. It lets the wizard select for its responses (partially) pre-defined utterances, enter query data and execute queries. Furthermore, it keeps track of the course of the dialogue by storing in a log file all utterances, their time stamps, and some other useful information, like whether wizard's utterances were selected from a menu, or self-invented and typed.

The screen of the Wizard Interface is composed of a number of windows: The most important ones are the wizard-subject communication window and the database interface window. The figures 1–5 and their captions show these windows and explain the operations they allow.

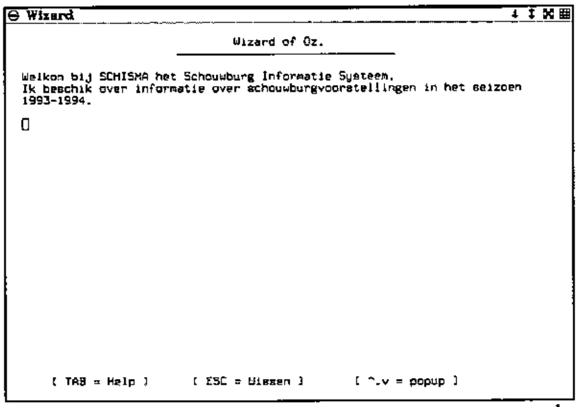

figure 1: initial wizard-subject communication window[1]

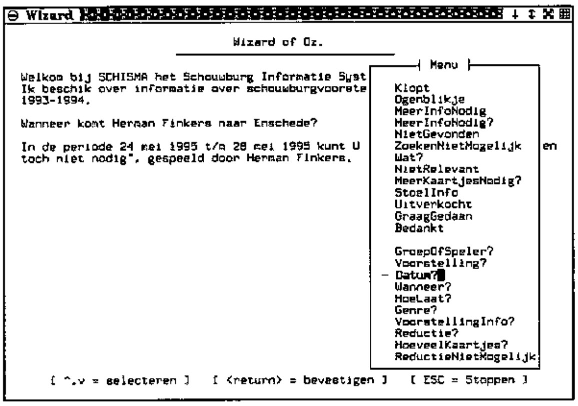

figure 2: wizard-subject communication window with pop-up menu of possible utterances.

The Subject Interface has one window, which looks a lot like the wizard-subject communication window the wizard has. In this window both the subject's own utterances and the wizard's appear. The wizard's utterances are displayed per utterance to the subject to obtain a computer-like way of presenting utterances. The subject does not have the option to select an utterance from a menu of canned sentences, as the wizard has (cf. figure 2).

All programs are written in C. The Subject and Wizard Interface both use the socket mechanism to exchange information.

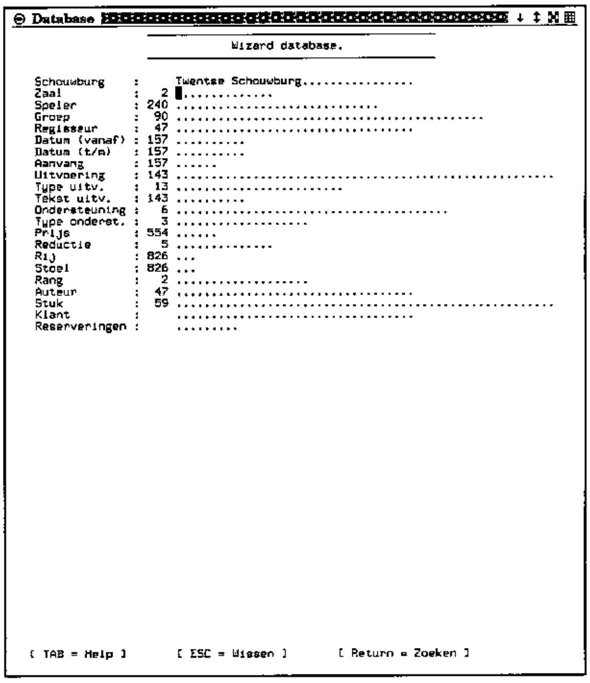

figure 3: the database interface window in its initial state.

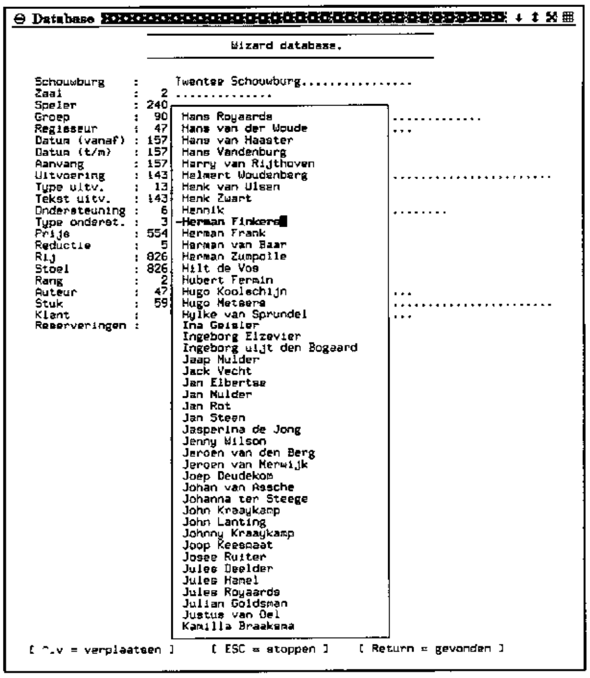

figure 4: the database interface with pop-up menu of possible actors.

[1]Figures 1 to 5 were produced by Jos Buis, who designed and implemented the wizard environment.

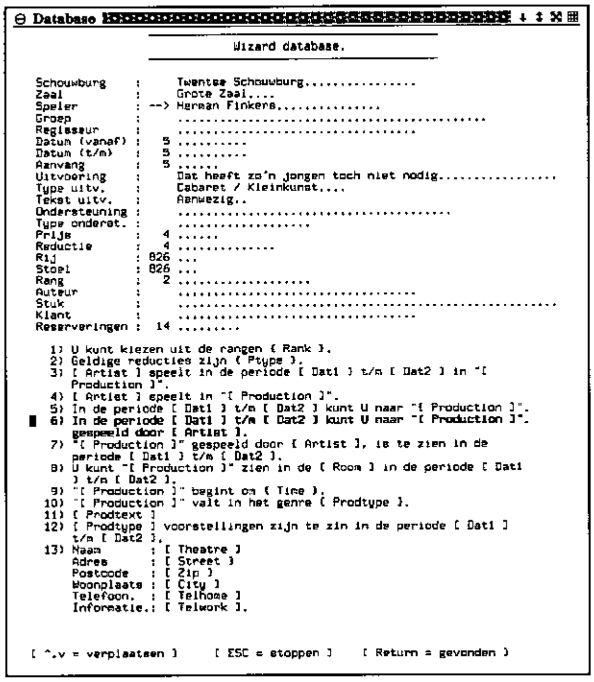

figure 5: the database interface window with templates for utterances the wizard can choose from.

In addition to its contribution to collecting a corpus of relevant dialogues, the Simulation Environment serves another goal in the project. It is implemented in such a way that it can easily be extended with components which take over some task of the wizard.

E.g. the wizard's task may switch from providing data on a fill-in form for database queries to providing data on a fill-in form for parse results. For this purpose, a component has been added to the environment that may turn the parse result into the intended query. This component could be one that is designed for the completely automatic system. Thus, the Simulation Environment is also a testbed for parts of the final system.

2.2. SET-UP OF EXPERIMENTS.

It will be clear that it is not an easy task to let experimental dialogues evolve smoothly. Every experiment will take 'tuning' of environment as well as wizard, wizard's assistant and instructor in a number of phases.

Also the order of experiments is crucial. Guyomard and Siroux explain in [11] how they conducted spoken Wizard of Oz experiments in two steps. In the first step of their experiments, data were gathered using a strict dialogue model and in the second phase those data were used for a more free dialogue model. In SCHISMA a different order is chosen. Insight in the more varied linguistic behaviour of subjects under a more free condition is considered a prerequisite for the design of dialogue models. So, first there will be an experiment in which no dialogue models are used and then an experiment in which a stricter dialogue model is used which is derived from the previous experiments.

Each experiment consists of the following phases:

The pre-experimental phase. In this phase, the experimental set-up was determined and the Simulation Environment was designed and implemented. The scenarios to be used in the pilot experiment were written, as well as the instruction sheets, which should both help and constrain the subject in his dialogue with the wizard.

The pre-experimental tests were carried out with the Computer Science Department of the University of Twente as the test site for the wizard and his assistant, and PTT Research as the test site for the subjects and their instructor.

The pilot phase. The pilot phase will be carried out at the University of Twente and several other universities. It serves the purposes of testing the Simulation Environment on the ease to use it, the consequent behaviour it exposes and on the kind of dialogues it yields, and letting the wizard get acquainted with the Simulation Environment.

The main experimental phase. In this phase, the actual dialogue sessions take place. These sessions will be carried out in a realistic environment in which the subjects are people who actually planned to inquire about a theatre performance or to make a reservation.

The analysis phase. In this phase the subject-wizard dialogues will be analyzed on several levels. The section below is entirely devoted to this subject.

3. THE ANALYSIS OF DIALOGUES.

In the analysis of dialogues we will distinguish three aspects.

1. Identification of the components a dialogue is built from (things like utterances and turns).
2. Tagging (the components of) dialogues.
3. Evaluating the annotated corpus, validating or rejecting hypotheses about dialogue structure, developing new hypotheses about dialogue structure.

In this section we shall not discuss the third aspect. We concentrate on aspects 1 and 2.

3.1. Dialogue components.

The first aspect of analysis is not really a separate analysis issue. It influences decisions that are made in implementing the wizard environment, and it is in turn influenced by such decisions.

We shall not dwell upon the decisions about what is an utterance and what is a turn. Let us simply state that we have taken a very pragmatic stand point here, relying on the fact that we deal with keyboard input. An utterance is anything that ends with pressing the return key. A turn is defined in terms of the time stamps associated to consecutive utterances. If no utterance of the other party is timed in between, then two utterances belong to the same turn. The wizard environment provides the timing information.

```
---Dialoog.Jos Buis.941118.03
---Version 0, created by the system.

{!client_age 20}
{!client_gender M}
{!client_assignment 3}
{!client_start_time 15:54:08}

C.1.1:
wanneer zijn er musicals?
{time 15:54:18}

S.2.1:
{generated}
Musical voorstellingen zijn te zien in de periode woensdag 28 december 1994 t/m vrijdag 19 mei 1995.
{time 15:54:46}
{continue 1}
{generated}
U kunt naar de volgende musical voorstellingen : Bubbling Brown Sugar, Candide, Little voice, Sweeney Todd, The marriage of Figaro en Under a blue Roof.
{time 15:55:19}

C.3.2:
wanneer is candide
{time 15:55:49}

three turns deleted

C.7.4:
reserveer vier kaartjes voor candide
{time 15:57:14}

S.8.4:
{generated}
De Grote Zaal heeft de volgende rangen : 1e balkon 1 en zaal 1. Welke rang wilt u zitten?
{time 15:57:28}

C.9.5:
zaal 1
{time 15:57:37}

S.10.5:
{generated}
Dit theater kent de volgende reducties : CJP, Geen, JTK en Pas 65. Heeft u een reductiekaart?
{time 15:57:46}

C.11.6:
nee
{time 15:57:52}

S.12.6:
{edited}
De totale kosten zijn f 104,00. Moet ik de kaartjes reserveren?
{time 15:58:28}

C.13.7:
ja
{time 15:58:34}

S.14.7:
{generated}
U heeft gereserveerd voor "Candide", gespeeld door Voltaire op woensdag 1 februari 1995. Aanvang van de voorstelling is 20:00. U wordt verzocht deze kaartjes minimaal een half uur voor aanvang van de voorstelling af te halen.
{time 15:58:39}

{!total_time 00:04:59}
{!total_turns 14}
```

figure 6: a dialogue as logged by the Wizard.

In figure 6 a part of a subject-wizard dialogue is shown. One may notice that turns are distinguished by a header line which signals who the 'speaker' was, namely S for system or C for client. The header also shows the turn count, in total as well as per speaker. A time stamp is given at the end of each utterance, if two or more utterances belong to the same turn there are additional *continue* labels. Every utterance of the wizard (system) is marked as either typed, or selected from a menu. The figure shows the contents of the text file exactly as it was logged by the wizard environment.

There are more components to be distinguished in a dialogue: there are turns and utterances, but also sentences, phrases, and words. It suffices here to say that our definition of these concepts must take into account that we deal with keyboard dialogues. A phrase which is just a single question mark (as a cry for help) must be considered as well.

3.2. DIALOGUE TAGGING.

Assigning annotations or tags to dialogues is important for other parts of the project. We mention three grounds for our interest in obtaining a tagged corpus of theatre information dialogues.

Firstly, without annotations it is impossible to move to the third aspect of analysis: developing dialogue models and testing hypotheses about dialogue structure.

Secondly, the annotated dialogues can serve as material to run tests of parts of the final system on.

Finally, some parts of the system will not run without data obtained from the final system. This is in particular the case for the discourse transition matrices that we discuss in section 5.

In this paper any tagging that has to do with the lower levels in the hierarchy of components, such as syntactic tagging, will be neglected. Such tagging is interesting and relevant, but the emphasis in this paper is on dialogue management, so here we look only at dialogue structure tagging.

The tagging related to dialogue structure is found mainly at the utterance and turn level. Some of these utterance and turn tags will be considered now in more detail.

One kind of dialogue structure tagging is derived from the topic-focus approach, and tries to capture something like *subject matter* of an utterance. The tags are highly domain dependent, there is a limited number of attributes which qualify as possible subject matter in a relevant dialogue. Things that count are *performance, place, date, time, number of seats, rank, price* and a few more. The tags are sets of such attributes. Dialogues with this subject matter annotation will be used in three ways.

Firstly, it is assumed that all dialogues will show a common progression in subject matter. Stated otherwise, the hypothesis is that there is a preferred order of dealing with the relevant aspects of the booking topic. The availability of a tagged corpus will help to gain insight on this hypothesis. We return to this preferred order in section 5 below. Note that we talk here about an ordering among subject matters which are all equally possible as next subject when viewing the 'information state' and the current 'theme' in the dialogue. Restrictions on the choice of a next subject that have to do with the status of available information are discussed in section 4.

Secondly, the final system must be capable of deriving the subject matter of an utterance, possibly using various sources of knowledge. The components of the system which should perform this task can be tested if an annotated corpus is available.

Finally, although this is dependent on the development of our view with respect to the 'preferred order of subject' hypothesis as formulated above, we envisage that counts of subject changes will appear in discourse transition matrices in the final system. This point is further elaborated in section 5.

Another kind of dialogue structure tagging is the initiative-response tagging. It has been pointed out by others (Bilange [5], Stubbs [19]) that every dialogue can be subdivided into initiative-response, or maybe initiative-response-confirmation units, and that such units could be viewed as the minimal constituents of a dialogue. Part of dialogue analysis is finding and marking these constituents. We do so by tagging at the turn level for initiative or response. This kind of tagging has been applied and extensively investigated by others also, in particular by Dahlbäck (cf. [1], [6] and [7]).

The objectives we pursue by providing this kind of tagging are largely the same as the ones we mentioned before. We return to the issues of initiative and response in section 5.

There are yet other kinds of tagging, at the turn and utterance level, that we consider, but we shall refrain from giving a complete list. In fact there is no complete list at the moment. A serious effort at tagging dialogues has started only recently, and the hypotheses about regularities and irregularities in dialogues, as well as the variety of possibilities for tagging and their relevance are subject of discussion.

4. THE INFORMATION CONTENT OF UTTERANCES AND THE INFORMATION STATE OF A DIALOGUE.

One could say that the remarks in the previous section about the analysis of dialogues are superficial in the sense that they deal with surface structure only. So far we have avoided any discussion on semantics, intentions, or plans. To put it bluntly: so far we have not tried to capture in any formal setting or model what the dialogues we consider are really about. This aspect of functionality and its formal model is the theme of this section.

In general, the functionality of a dialogue is described as a series of transitions with respect to a finite set of possible states (cf. [2,5,14,21]). Each state corresponds to some intentional state of the actors: being (dis)satisfied by an answer, having a question, believing something not known by the other actor to be true, etc. According to this approach, the purpose of a dialogue will be the fulfilment of the intentions of the actors raising and raised by the discourse.

In our experimental system for theatre information and booking, the continuation of the dialogue also depends on the information state resulting from the most recent dialogue act. The identification process of a performance is less straightforward than a sub dialogue identifying a train or a flight. Being aware of the vagueness and indistinctness of the keywords related to this domain, users utter additional information in order to check whether they were right in their expectations. For example, instead of using the identifying phrase *the performance next Tuesday* or *the premiere of The Tempest*, users utter *the* ***opera*** *next Tuesday* or *the premiere of* ***Shakespeare's*** *Tempest.*

Not only the user, however, but also the system is aware of the vagueness of some technical terms. For instance, when selecting the seats the user would desire, the system first asks in what part of the theatre the seats are to be located and consequently presents a proposal. It turns out that in this particular domain it is far less clear when the selection criteria to be given by the user have already identified a unique performance or not and that appropriate action has to be performed as soon as certain criteria do exclude each other.

This analysis leads to an approach towards the functionality of dialogues in which information states play a more important role than in the approaches generally based on Searle's Speech Acts theory ([17]). In this theory information states have been simplified to the felicity condition that with respect to that particular information state the current speech act may be performed. Only a change of this condition, not a change of the information state, is represented and plays a role in determining the sequel of the dialogue.

It will be shown that the role information states play cannot simply be added to a speech acts based approach. In order to represent the role of information states, we must use information state changes as the foundation of our approach and regard the purpose of a dialogue to be the transfer of information. Therefore, the functionality of discourses, from the point of view of their objective, can be described in terms of conveying information. This change of approach gives rise to at least the following questions:

1. what is information and what exactly is the difference from intentional states?
2. in what way is information conveyed?
3. how is such information dealt with?

These questions will be subject of the following subsections.

4.1. WHAT IS INFORMATION?

At first sight it seems to be very odd that the term 'information' can be used both for *secure knowledge* (common use) as well as for an *increase of certainty* but still related to some probability and uncertainty (technical use). Both uses, however, are based on the common notion that information is knowledge to act upon. Knowledge, we are not entirely sure of that it is true, but which we assume to be so.

In this sense, Groenendijk et al [10] are right in perceiving the increase of information both as an increase of things we have to know something about as well as a decrease of possible state of things by knowing more about those things.

The contemporary use of the term 'information' originates with the American philosopher C.S. Peirce (1839–1914) ([15]). Confronted with the Kantian rule that the extension and intension of a term are reciprocal (that is, the more attributes a term has got, the less objects it refers to, and vice versa), Peirce states that according to that rule we are not able to learn something new.

According to him, we must be able to increase the number of attributes (depth), with respect to the same number of objects (breadth), while, on the other hand, it must also be possible to increase the breadth the depth remaining unchanged. The result of both processes will be an increase of information. And so, 'information' is defined by Peirce as *the amount of comprehension (=intension) a symbol has over and above what limits its extension.* That is to say, information, as defined by Peirce, consists of attributes which are additional with respect to the attributes needed in order to determine the selected data. Combined with the Peircean formula

Extension × Comprehension = Information

expressing his rejection of the Kantian formula

Extension × Comprehension = Constant

according to which indeed the attributes contained by a predicate will be represented by the selected extension properly, we have to represent information as a tuple containing both a set of attributes and a set of entities the attributes may be predicated to.

4.2. IN WHAT WAY IS INFORMATION CONVEYED?

In order to get a clear view on the matter, one needs to distinguish between the manners in which information can be conveyed, and the way in which such information is actually managed. Let us first attend the issue of the manners in which information can be conveyed.

Central to our treatment of the communication of information are the familiar notions of *topic* and *focus*, as developed by the Prague School of Linguistics. Topic, in general, signifies the already available, or 'given', information, while focus expresses additional, or 'new', information.

Due to the articulation of topic and focus in a sentence, and the relations between topics and foci of different sentences, a certain story-line or thematic progression is developed in the discourse.

Four kinds of thematic progression are distinguished by Daneš[2]:

1. Sequential progression: $T_1 \rightarrow F_1$, $T_2 \rightarrow F_2$, ...; T_2 is constituted by F_1
2. Parallel progression: $T_1 \rightarrow F_1$, $T_2 \rightarrow F_2$, ...; $T_1 \approx T_2$, i.e. are highly similar
3. Hyper theme: $T_1 \rightarrow F_1$, $T_2 \rightarrow F_2$, ...; T_1, T_2 refer to hyper theme T_H

The interesting fact of defining the kinds of thematic progression as such, is that we arrive at building blocks for the communication of information. Each block is characterized by conditions, induced information change, and a relation between the sentences. The relation indicates whether the information is negative or positive, while the conditions state the applicability of the building block – i.e. whether a certain kind of information can actually, sensibly, be communicated in a certain situation. This is outlined in some more detail in [13].

4.3. HOW IS SUCH INFORMATION DEALT WITH?

As already stated above, the dynamics of interpretation can be perceived to lay in the change of information conveyed by the discourse. In their theory of Dynamic Semantics, which is currently under development, Groenendijk et al. [10] emphasize that the dynamics of interpretation should be reflected in the formal representation. In particular, meaning is to be understood as information change potential. According to Groenendijk et al, the communicated information is strictly positive. Hence, each utterance in the discourse is conceived of as an update of the hearer's information. Revisions or downdates are not allowed, which might be marked as a first shortcoming of Dynamic Semantics. The formal representation of Dynamic Semantics, for which we refer to Groenendijk et al. [10], is characterized by the use of possibilities making up information states.

The possibilities intuitively stand for those alternatives open according to the present information. Formally, possibilities consist of pegs each of which is associated to an object in some possible world. A second shortcoming might thus be noted. The predicates, of which mentioned objects are the subjects, are not represented in the possibilities. Holding on to the original ideas of Groenendijk et al, we have proposed a revision of the representation of Dynamic Semantics that is based on Peirce's logical theory of information ([16]). For our revised representation we define a similar notion, namely that of

[2] The fourth kind, Prosody is not defined very clearly by Daneš and hence it will not be treated in the sequel of this section.

the spot, consisting of a set of attributes and a set of entities the attributes are predicated to.

Each predicate has an epistemological modal attached to it, where ◊ is a *may-be* or *possibility*, O an *is* or *actuality*, and □ a *should-be* or *tendentiality*. By these modals we are enabled to distinguish the essential and additional predicates as well as the constraints related to the domain.

Rather than being an index to an entity perceived actual in a possible world, the spot is a symbol, that matures by comprehending more accurate information about the world discoursed of. By 'more accurate information' we understand information states arrived at by using a combination of updates and downdates, depending on the discourse. Hence, discourse is perceived of as a process of conveying information being either positive or negative. Recall that Groenendijk et al state that discourse is a process of conveying strictly positive information. Our perception leads to a series of definitions concerning information states, updates, downdates, revisions, etc. which is rather similar to, though more extensive than, Groenendijk's.

5. THE GUIDING ROLE OF DISCOURSE TRANSITION MATRICES.

In the previous section we talked about the functionality of a dialogue in terms of a set of possible states, and transitions between them. It was pointed out that state in SCHISMA dialogues is largely information state. Ways of representing such states, with emphasis on the state change potential, were proposed and discussed.

In this section we leave the functionality aspect and concentrate more on form. The main observations of this section are the following.

Firstly, state in our dialogues may be, as we saw, largely information state, but it is not exclusively so; there are other relevant aspects to the notion of current state.

Secondly, it is worthwhile to consider and represent the other aspects of state independently of the more important information aspect; determining and effectuating state transitions is in our view the process of determining the optimal combination of two or more independent choices.

So we return to states and transitions, but we look at other aspects and, for the moment, ignore the aspect of information state and conveying information.

One could say that what we talk about here are the protocols that play a role in conducting a, or even any, dialogue. Things that fall under this heading are:

the natural sequences of initiative and response and the disruption of such sequences,

the choice of subject and the ordering of subjects to be discussed,

etc.

One can imagine how large parts of a dialogue with SCHISMA consist of simple question answer pairs, where the initiative (question) lies with the user, and the system's behaviour is purely responsive (answering). But once it has become clear for what performance on what day the user wants a booking, the initiative and response roles may interchange. Now it can be the system questioning the user about his preference for seats in the theatre, or for his possible rights on a reduced price. There are other points in a conversation where the initiative may change from one partner to the other. E.g. questions by the system about preference for seats could well be countered by the client posing a question about prices first. The other way round, an open question by the client about performances this season could lead to a counter question by the system, urging the user to be more specific. The same may happen if the client poses a question which is too vague, or one that is in some way inconsistent in itself.

It is essential that the system is constantly aware of which partner holds the initiative in the dialogue, and how the initiative changes. This will help understanding as well as generating utterances. It may also help to prevent dialogues going wrong because both partners start claiming the initiative, or both partners keep refusing to take the initiative.

Moreover, being aware of who has the initiative is important for a consistent behaviour of the system. Such consistency can be realized in different ways. One way to be consistent is to go for the initiative as much as possible. But it is also consistent to leave the initiative with the user as much as possible.

This is not the place to discuss the proper choice between the two (or may be even more) behaviours. The point is, that initiative is a factor in conducting a dialogue. Initiative is related to the conveying of information, but initiative is not part of the 'information state'.

Another aspect of consistent behaviour, which is has to do with conveying information, but which is a matter of protocol rather than being a part of an information state, is the ordering of subject matters (foci) in a conversation. We mentioned this point already in our discussion on the tagging of dialogues in section 3.

An example of a situation where this ordering seems to be relevant is, when in a dialogue the client has established a date and a performance (s)he wants to go to, and the initiative is left for the system to guide the client through the actual booking process. Although part of the booking will certainly be to inform the client about the price (s)he has to pay for the tickets, and although this price may very well depend on whether the client is entitled to some sort of reduction, it seems extremely awkward to open the booking process with a question about the user's claims for such a reduction. The number of persons that want a ticket, and preference for a rank or a particular position of the seats in the theatre seem far more natural first topics.

Apparently a smooth dialogue needs a smooth ordering of issues to be discussed.

The system must be constantly aware of the most natural (or most probable) next subject to come into the dialogue. This will contribute to an easier understanding of utterances by the client, as well as to a better choice of a question or a remark to be generated, if the initiative is at the system's side.

The initiative in a dialogue and the subject-matter of utterances that we discuss here, were also a theme in the section on dialogue tagging. The point of what was said in this section so far has been that being aware of various kinds of states and being consistent in state transitions is important for smooth dialogues.

In the concluding remarks we will now present a link between the tagging, dialogue states and transitions between them. This link should at the same time explain the title of this section.

It is our idea to equip the final system with one or more transition matrices based on n-gram counts of tagging as found in our annotated corpus. The 'protocol-state' one could say is the history of the dialogue restricted to its last n events. The transition matrix gives the likelihood for every next event in every state.

The data needed to fill the matrix are derived from the numbers as found by counting occurrences of states in the corpus. Note however that we are not obliged to follow these numbers exactly. Consistent behaviour could be improved by tuning some values by hand.

In any case, the transition matrix or matrices are simply Markov models for aspects of the protocol in SCHISMA dialogues!

It must be re-emphasized however, that state as we discuss it here is of restricted importance. There is information state as well, and the final state transition of the dialogue system will be chosen primarily on the basis of that state. The protocol state helps, it should make the flow of information and things like the changes of roles between dialogue partners, the awareness of mental state (like disappointment or discontentness) go smoother.

6. USING NATURAL LANGUAGE.

In SCHISMA, we put quite some effort in the development of a natural language interface using a commercial tool set for building such interfaces: the Intelligent CONnector (ICON, also known as Natural Language) from Natural Language Inc.

The reasons to build such an interface were:

- to evaluate this tool set which according to Sijtsma and Zweekhorst [18], is one of the most sophisticated;
- to investigate to which extent a dialogue can be handled by question answering systems like this one;
- to have a question answering system that could function in a comparison of results of the SCHISMA project with a similar functionality.

Although the final version and the evaluation report will be ready by the end of 1994 (Komen [12]), some preliminary observations can be reported already.

Firstly, our expectations about the ease of use of the tool set and the quality of the resulting system diminished during the development. While 70% of interface was built in two or three weeks, the next ten to fifteen percent was a lot harder to accomplish. At 85% percent the process of steady progress changed into the contrary: with each new addition, the system could collapse to a state in which it misinterpreted questions that it answered correctly earlier.

Secondly, the resulting system is clearly not of the kind that could function as a co-operative partner in a quest for information. The interpretation which is the basis for the transformation of a question to SQL, treats questions (almost) completely isolated from its history. (The only exceptions are anaphoric expression such as *it* or *he*.) The system cannot handle open, explorative questions or statements like *I'd like to go to an opera tonight* and it is not possible to stick to a particular subject for a while. Furthermore it is impossible to model the effects of reservations such as *Please, reserve two seats at the performance tonight.*

As a last remark we would like to mention the (sometimes extreme) bias towards financial and administrative domains.

Although we still think it valuable to have comparable systems at hand, we can already conclude that any dialogue system will outperform systems built with ICON, simply because they cannot handle dialogues at all.

7. TOWARDS SCHISMA 2001.

As explained in the previous sections, different research lines are pursued in the SCHISMA project. As announced in the introduction, we shall use this final section to provide a broader perspective on things to come.

It is the intention of the Parlevink research group in Twente, that also in the future SCHISMA will allow and stimulate different and changing goals. SCHISMA should be a vehicle in which research results of researchers and Ph.D. students will be embedded.

This does not mean that there will be no end products or tools that can be offered to interested partners. Such tools are available already at this moment (maybe not as a fully developed product, but certainly usable).

To mention a few, there are:

- a system to get information about theatre performances which is developed using Natural Language;
- a Wizard of Oz tool, which may find application in quite different domains;
- an environment which allows investigation of feature structures in a user-friendly way;
- a PROLOG implementation of the information state model as described in section 4.

However, in addition to current activities, many research directions, with further reaching technical aims, will be followed. Below some of them are listed.

Speech. Obviously, the most natural extension to SCHISMA is making it accessible by speech, preferably by telephone. It does not seem difficult to extend our Wizard of Oz tool in such a way that a spoken language corpus for the theatre domain can be obtained. More knowledge about spoken language access to a kind of SCHISMA system will be obtained from interviews that have been and will be conducted at some theatres in some Dutch cities. We have not planned yet to record real-world dialogues by telephone of theatre information or booking services. It is hoped that this can be done in a similar way as is being planned in the framework of the Dutch "Prioriteitsprogramma Taal en Spraak" for train travel information. Presently we plan to investigate speech especially with the aim of integrating speech and language. But there are options for more extensive neural network research into speech. They are discussed below.

Integrating speech and language. Traditional speech and language research have led to different research communities. Speech research is very much concerned with signal processing, noise filtering, phoneme and word recognition. The use of higher-level resources (syntax and semantics) is seldomly employed. The use of statistical approaches is advocated. In SCHISMA research corpus-based and statistical methods will play an important role. Apart from that, we expect it to be possible to integrate our methods of syntactic analysis and unification with HMM speech recognition. The same holds true for knowledge of topic-focus articulation, prosodic information and integration of these aspects in a parser.

Robustness. It is well known that in the near future no comprehensive formal and effective models for all kinds of knowledge sources that play a role in language interpretation will be available. Clearly, in this case analysis with respect to a certain knowledge source can not always lead to a desired result. More in general, if the input is not according to the model of the knowledge source it cannot be given a correct interpretation with respect to this knowledge. Since a complete analysis with respect to one particular kind of linguistic or extra-linguistic knowledge will never be goal in practical applications it is useful to see whether incomplete analysis with respect to a certain type of knowledge can be compensated by using a different type of knowledge. This compensation can be considered as adding robustness to a certain level (or viewpoint) of analysis (and therefore to the whole trajectory of analysis). For example, the use of unknown words may be compensated by syntactic knowledge, but also by knowledge about punctuation or the way humans make errors in typing or speaking. Similarly, not allowable syntactic constructs can be repaired using certain semantic knowledge. Research on robustness, also including probabilistic methods will remain part of the project and the results will be incorporated into (the) system(s) that will be built.

Evaluation issues. Measuring the quality with which software performs a certain task is an issue that has also emerged in the field of natural language processing. Clearly, due to incomplete models of nearly all aspects of natural language evaluation is a difficult question. Evaluation in the SCHISMA context, for example, can mean evaluation of the system as a whole, evaluation of the different parts or evaluation with respect to different groups of users. Evaluation standards are hardly available but we expect that in the near future the attention that is going to this subject (e.g., the ARPA speech contests, the MUC contest, the morpholympics, the parseolympics) will pay off. Although there is no real competition we hope that in the future at least comparisons can be made between parser systems for Dutch (head and left corner (unification) parsers, data-oriented parsers, AGFL parsers, etc.) and natural language interfaces for Dutch (the primitive SCHISMA obtained from the commercial 'Natural Language' tool, SCHISMA under development and the OVIS system that will be built in the framework of the Dutch "Prioriteitsprogramma Taal en Spraak"). For literature on these issues see Thompson [20].

Neural and Fuzzy Approaches. Probably the most straightforward application of neural network techniques in SCHISMA is in speech recognition. Although some modest excursions into this area have been made (e.g., M.Sc. research by students) there is insufficient expertise and manpower to enter this field. A joint research proposal with, among others, the Max Planck Institute of Nijmegen, will hopefully give us the opportunity to make a start in this field and to use results in SCHISMA. For a more comprehensive approach to neural network techniques for speech and language that deserves consideration for use in the SCHISMA context see Drossaers ([9],and also this volume). Presently it is far from clear whether complete integrated linguistic analysis with neural networks for the SCHISMA domain can ever be realized. We nevertheless assume that the sequence recognition capabilities of the network that has been developed can support decision making in several sub tasks (robustness, disambiguation, dialogue modelling, etc.) for SCHISMA. In addition to the neural network research some modest efforts are being made to develop language theory using a fuzzy logic approach.

At this moment there are no plans to make SCHISMA multi-modal. The emphasis is on keyboard input and if possible we would like to add the possibility to access SCHISMA using speech input. In order to continue SCHISMA research in a satisfactory way it will be necessary to strengthen the cooperation with PTT Research and to embed SCHISMA research in several Dutch and European research programs.

LITERATURE.

[1] Ahrenberg, L., A. Jönsson, and N. Dahlbäck. 1990. Discourse representation and discourse management. In: Dahl and Fraurud (eds.) *Papers from the 2nd Nordic Conference on Text Comprehension in Man and Machine*, Institute of Linguistics, Stockholm.

[2] Allen, J.F. and C.R. Perrault. 1980. Analyzing Intention in Utterances. *Artificial Intelligence 15.*

[3] Andernach, T., G. Deville, and L. Mortier. 1993. The design of a real world wizard of oz experiment for a speech driven telephone directory information system. In: *Proceedings of Eurospeech.*

[4] Andry, F., E. Bilange, F. Charpentier, K. Choukri, M. Ponamale, and S. Soudoplatoff. 1990. Computerised simulation tools for the design of an oral dialogue system.

[5] Bilange, E. 1991. A Task Independent Oral Dialogue Model. *Proceedings of the 5th Conference of the European Chapter of the Association for Computational Linguistics.*

[6] Dahlbäck, N. and A. Jönsson. 1986. A system for studying human-computer dialogues in natural language. Research report, NLPLAB IDA Linköping University, Linköping, Sweden.

[7] Dahlbäck, N. and A. Jönsson. 1989. Empirical studies of discourse representation for natural language Interfaces. In: *Proceedings of the 4th Conference of the EACL.*

[8] Daneš, F. 1974. Functional sentence perspective and the organization of text. In: F. Daneš (ed.), *Papers on Functional Sentence Perspective.* Prague: Academia.

[9] Drossaers, M.F.J. 1995. *Neural Networks for Integrated Linguistic Analysis.* Ph.D. Thesis, Parlevink Research Group, University of Twente, (to appear).

[10] Groenendijk J., M. Stokhof, and F. Veltman. 1994. Dynamic Semantics. *Course Material Sixth European Summer school on Logic, Language and Information.*

[11] Guyomard, M. and J. Siroux. 1988. Experimentation in the specification of an oral dialogue. In: H. Niemann, M. Lang, and G. Sagerer (eds.), *Recent Advances in Speech Understanding and Dialog Systems.* Springer Verlag, Berlin.

[12] Komen, E. 1994. Building SCHISMA in Natural Language. *Memoranda Informatica.* University of Twente, Enschede. (to appear)

[13] Kruijff G-J.M. and J. Schaake. 1994. Discerning Relevant Information in Discourses Using TFA. Presentation at the Fifth Symposium on Computational Linguistics In the Netherlands CLIN94.

[14] Lochbaum, K.E. 1991. An Algorithm for Plan Recognition in Collaborative Discourse. *Proceedings of the 29th Annual Meeting of the Association for Computational Linguistics.*

[15] Peirce, C.S. 1982. *Writings of Charles S. Peirce: a Chronological Edition.* Edited by M. Fisch et al. Vol. 1: 1857-1866. Indiana University Press, Bloomington.

[16] Schaake, J. and G-J.M. Kruijff. 1994. Information States Based Analysis of Dialogues. Presentation at the Fifth Symposium on Computational Linguistics In the Netherlands CLIN94. 1994

[17] Searle, J.R. 1969. *Speech Acts, An Essay in the Philosophy of Language.* Cambridge University Press, Cambridge.

[18] Sijtsma, W. and O. Zweekhorst. 1993. Comparison and Review of Commercial Natural Language Interfaces. In: F.M.G. de Jong and A. Nijholt (eds.), *Natural Language Interfaces: From Laboratory to Commercial User Environment. Proceedings of TWLT5.* University of Twente, Enschede.

[19] Stubbs, M. 1983. *Discourse Analysis.* Blackwell, Oxford, 1983.

[20] Thompson, H. (ed.). The Strategic Role of Evaluation in NLP. Proceedings, Edinburgh, Scotland, April-May 1992.

[21] Wachtel, T. 1986. Pragmatic Sensitivity in NL Interfaces and the Structure of Conversation. *Proceedings of the 11th International Conference on Computational Linguistics.*