ATNS USED AS A PROCEDURAL DIALOG MODEL

D. Metzing, University of Bielefeld,
Germany

Summary

An attempt has been made to use an Augmented Transition Network[10] as a 'procedural dialog model'. The development of such a model appears to be important in several respects:

- as a device to represent and to use different dialog schemata proposed in empirical conversation analysis;
- as a device to represent and to use models of verbal interaction;
- as a device combining knowledge about dialog schemata and about verbal interaction with knowledge about task-oriented and goal-directed dialogs.

A standard ATN should be further developed in order to account for the verbal interactions of task-oriented dialogs.

Introduction

Naturally occuring task-oriented dialogs[4] are the joint product of the interactions of (at least) two participants who know how to *cooperate*, i. e. who know how to organize

- their social interactions
- their verbal interactions, and
- their task-oriented interactions.

The *amount of such interactions* which seem to be necessary in specific task-oriented dialogs may

(i) depend on a number of factors given in advance such as: readiness to cooperate (a), preciseness of the task representation (b), amount of mutually shared task-specific knowledge (c), amount of knowledge about the other participant (d) and personal factors as for example competence and (self-) confidence (e);

(ii) depend on procedures apt to modify these factors (a - e) in an efficient and positive way;

(iii) depend on procedures used for task resolution and result explanation.

Participants of naturally occuring task-oriented dialogs are able to make use of these factors and procedures in a skillful and flexible way, but such properties are still lacking even in experimental dialog systems.

In past natural language processing research considerable efforts have been made to process the structures underlying sentences or texts. Procedures have been developed which build up deep structures of sentences or which determine macro-structures or event skripts ('frames') underlying texts. In the next two years special efforts will be made to process the structures underlying task-oriented dialogs.

Representation of Interactional Knowledge

In coversation analysis, systematic accounts of the sequential organization of dialog interactions have been developed, e. g. for turn taking, for opening sequences, for closing or repair sequences[7] or for different types of task-oriented verbal interaction as a whole (e. g. giving advice, directions, explanations)[12]. But these accounts have only been of a *structural* type, not of a *procedural* type. A formal representation has rarely been attempted[6] and an integration or interaction of different knowledge sources is generally not considered.

In the subsequent sections we will argue for a *level of representation guiding the social interactional and the verbal interactional aspects of task-oriented dialogs*. A personal belief or knowledge component will use information of this *interactional* level together with information of a task level as well as information of a sentence/text level. We will argue for a procedural representation of interactional knowledge and we think that the usefulness of ATNs[10,3] for such a representation should be examined in more detail.

'Parsing Interactions'

The approach presented here differs from other computational dialog models in the following way:

- A dialog model is not based on an underlying dialog prototype[2] specifying essentially task-oriented information. It is claimed that the social interactional and the verbal interactional aspects of task-oriented dialogs are important enough to be represented in a detailed way on a special level.
- Dialog proporties are not only examined by problem solving techniques[1]. Instead, extended parsing techniques are used in 'parsing interactions' and they are supposed to be helpful in determining the interactional structure underlying utterances.

Let us further specify the kind of interactional knowledge which participants of certain types of task-oriented dialogs are supposed to have as well as ways to represent it. The participants will generally know how to manage the social, the verbal and the task-oriented interactions. They will generally know about several rather invariant components of a certain type of task-oriented dialogs as well as of a normal sequence of these components. They know about the detailed (alternative) structures of each component, the choice of which may depend on factors as were mentioned above(a - e). They know

- how to initiate a social contact/a verbal interaction and how to respond positively/negatively to this initiative;
- how to continue/to maintain an interaction,
- how to signalize interest, competence or (difficulties of) understanding,
- how to organize turn-taking,
- how to initiate the termination of a social contact/a verbal interaction and how to respond positively or negatively to it.

Part of this knowledge may be described as sequences of social/verbal interactions, formally to be represented as connected (sub-) networks of an ATN. I. e. social/verbal interaction is seen as a process of path selection of (at least) two participants in a network of states and state transitions. These (sub-) networks should be set up on an empirical basis (recordings of naturally occuring task-oriented dialogs).

An Example

Some properties of the interactional knowledge mentioned above may be represented in an ATN in a straightforward way, whereas the representation of others seem less straightforward.

Constituents

Let us assume that some types of information-giving dialogs may be assigned a

network structure with the following constituents of social/verbal interaction:

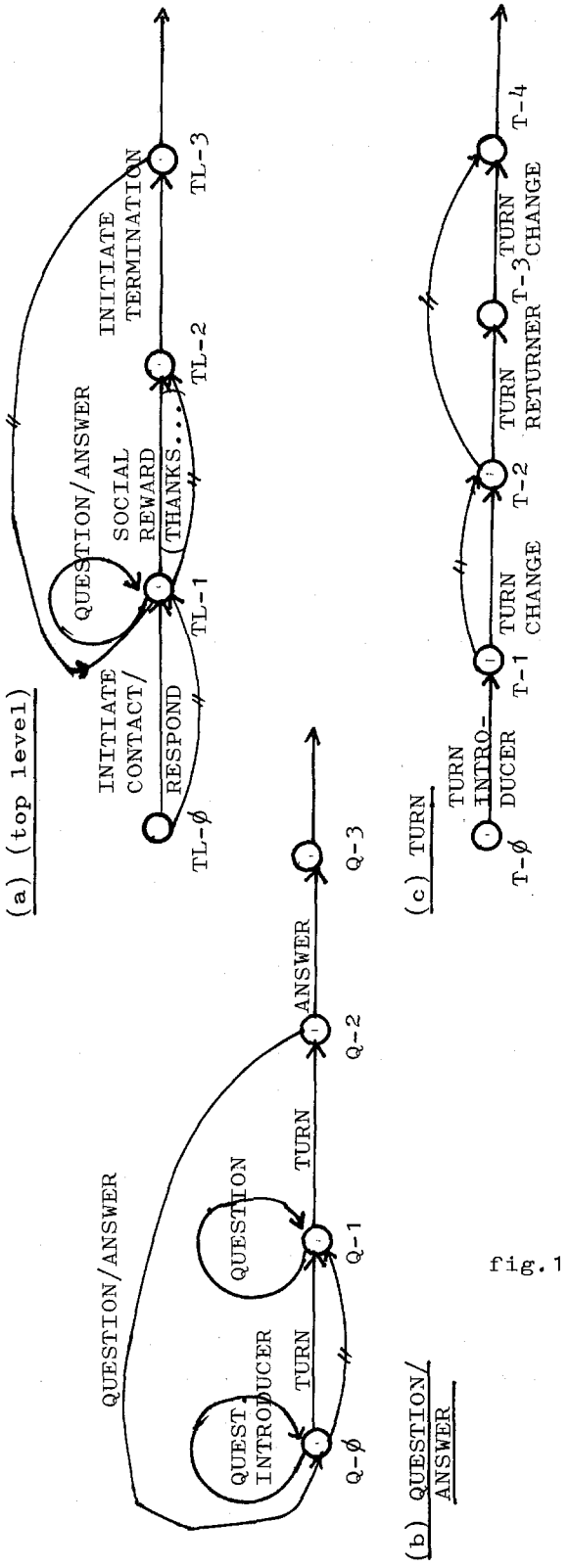

fig.1

Each (sub-)network represents sequences of interaction
- on the level of a dialog type (a),
- on the level of speech act sequences (b), or
- on the level of turn-taking organization (c).

The arcs indicate (alternative) state transitions/(alternative) courses of interaction. They may be marked or unmarked ('JUMPs').

According to fig. 1 a dialog interaction may start:
- with a contact initiating utterance ("EXCUSE ME ..."), level a;
- with a question introducer ("MAY I ASK YOU ...") level b;
- with a task-specific question ("WHERE DO I FIND ..."), level b;
- with a 'turn introducer' ("YES", "..."), level c.

Possible actions may be skipped (cf ⟶// arrow), repeated (loops; iteration) or some sequences may be embedded into other sequences (cf. QUESTION/ANSWER arc in the QUESTION/ANSWER subnetwork; recursion).

These subnetworks are connected to other networks with the same type of information or with other types of information:
linguistic information in order to recognize/generate different forms (e. g. to initiate a contact, to introduce questions, to pass turns) or direct/indirect ways to ask a question ("WHERE DO I FIND ..." vs "I DO NOT KNOW ...", "I SUPPOSE YOU KNOW ..."); *task-oriented* information in order to build up coherent sets of answers (on the basis of a task-model).

Note that some utterances may serve

several interactional roles ("MAY I ASK YOU ..." uttered at the beginning of a dialog initiates a contact *and* introduces a question.) Special tests on arcs will recognize this and the corresponding actions will build up an interactional structure according to such multiple roles of constituents. (A way to see them from different perspectives.)

Interpretation of Arcs

The interactional information represented in an ATN-subnet may be used to plan and to guide the recognition/generation of social/verbal interaction as part of a task-oriented dialog. The information represented on an arc should be used for both, for recognition as well as for planning/generation. The structure of the (sub-)networks is useful in the sense that 'normal' courses of interaction are explicity represented. So they are expected and indicate a kind of interactional coherence of a task-oriented dialog. But speakers may violate such normally respected dialog sequences and they can cope with this fact. It is therefore desirable to make a more flexible use of the information represented in a network. E. g. it seems desirable to calculate a (not yet existing) transition on the basis of task-specific cues and/or utterance cues (for example when a participant reopens an already closed subtask or when he repeats or reopens an already executed move or when he suddenly terminates an interaction.

Further developments

It seems desirable to have an extended ATN parser in order to cope with unexpected dialog sequences. In implementing aspects of social/verbal interactions one should examine carefully efforts made

- to use an ATN in a more flexible way[5,9]
- to combine recognition and generation in an ATN[8], and
- to build up several interacting ATNs".

References

1. Appelt, D. E. 1980 "Problem solving Applied to Language Generation", Proc. of the 18th Meeting of ACL, 59 - 63.
2. Bobrow, D. G., 1977 and the PARC understander group, "GUS - 1, a Frame Driven Dialog System" Artificial Intelligence 8 : 2.
3. Goldstein, I. P., 1976 M. L. Miller PAZATN: A Linguistic Approach to Automatic Analysis of Elementary Programming Protocols, MIT AI MEMO 388.
4. Grosz, B. 1978 "Discourse Knowledge", in: D. Walker (ed.) Understanding Spoken Language, 229 - 344
5. Kwasny, S., 1979 N. Sondheimer "Ungrammaticality and Extra-Grammaticality in Natural Language Understanding Systems", Proc. of the 17th Meeting of ACL, 19 - 24.
6. Metzing, D. 1979 Implementation of Interactional Processes Memo University of Bielefeld.
7. Sacks, U., E. Schegloff, G. Jefferson "A Simplest Systematics for Conversation

for the Organization
of Turn-Taking for
Conversation",
Language 50, 696-735.

8. Shapiro, S. C. "Generalized Augmented
1979 Transition Network
Grammars for Generation
from Semantic Networks",
Proc. of the 17th
Meeting of ACL, 25 - 29.

9. Weischedel, R. M., J. E. Black "If the
1980 Parser Fails"
Proc. of the 18th
Meeting of ACL,
95 - 96.

10. Woods, W..A. "Transition Network
1970 Grammars for Natural
Language Analysis",
CACM 13, 591 - 606.

11. - Research in Natural
1978 Language Understanding,
BBN-Report No. 3963

12. Wunderlich, D. "Ein Sequenzmuster
für Ratschläge: Ana-
lyse eines Beispiels",
to appear in: D. Metzing
(ed) Dialogmuster und
Dialogprozesse (Hamburg).