The Monroe Corpus

Amanda J. Stent

The University of Rochester
Computer Science Department
Rochester, New York 14627

Technical Report 728/TN 99-2

March 2000

Abstract

In this report we discuss the creation and initial annotation of the Monroe corpus , a collection of video and audio data of 20 human-human, mixed-initiative, task-oriented dialogs about disaster-handling tasks. We describe how the dialogs were collected, what tasks were used, and how the data was transcribed and aligned.[1]

This work was supported by ONR research grant N00014-95-1-1088, U.S. Air Force/Rome Labs research contract no. F30602-95-1-0025, NSF research grant no. IRI-9623665 and Columbia University/NSF research grant no. OPG: 1307.

[1]Our thanks go to George Ferguson for help with setting up hardware and for constructing the map, and to Peter Heeman and Mark Core for answering many questions about the speech-processing tools and DAMSL. We owe thanks to Kim Guenther and Leah Wasserman, two dedicated annotators. We also wish to thank the members of the NL group at the University of Rochester who helped refine the tasks and check the dialogs, especially Lucian Galescu, Donna Byron and Myrosia Dzikovska. Finally, we wish to thank James Allen for invaluable advice and assistance at every stage of this ongoing research.

A	Okay so there's ah *Cruel Intentions* at Hen- at Regal and
B	Mm-hm.
A	*200 Cigarettes* at Marketplace. [pause] Yes that's all.
B	Okay. What is *Cruel Intentions* rated?
A	Um [pause] R.
B	Alright. How about *Cruel* ...
A	[interrupts] Oh also *Godfather III* at Dryden.

A	*Cruel Intentions* is showing at Regal cinema and
	200 Cigarettes is showing at Marketplace cinema.
B	[pause] What is *Cruel Intentions* rated?
A	[pause] R.
B	[pause] Let's see *Cruel Intentions*.
A	[pause] *Godfather III* is also showing, at Dryden cinema.

Figure 1: Human-human and human-computer versions of a dialog

1 Introduction

Content-planning for dialog is a complex process in which an agent must plan not only to communicate its thoughts, but also to maintain the collaboration. Existing dialog systems do not produce the whole range of acts that can be observed in equivalent human-human dialogs. Consider the two constructed dialog extracts about selecting a movie in figure 1. The first is the kind of interaction two humans might have; the second is the same discussion as it might occur between a state-of-the-art dialog system and a well-trained human. While both interactions are successful, the second is more awkward and takes longer due to pauses and lack of grounding.

We are developing a theory of content planning that accounts for all the types of acts that may be produced in natural, mixed-initiative, task-oriented dialog: turn-taking acts, grounding acts, illocutionary acts and argumentation acts [22]. Our goal is the implementation of this theory in a conversational agent that could be an equal participant in such dialogs.

A natural first step in this process is the examination of appropriate human-human dialogs. We collected a corpus of mixed-initiative human-human dialogs in a complex domain, the Monroe corpus. We are annotating it for turn-taking behavior, speech acts and grounding acts, and rhetorical structure. This allows us to determine the relative frequencies of different types of acts, how they relate to each other and how they are combined in dialog contributions.

This paper describes the collection, transcription and alignment of the data. It also describes the annotation of eight of the dialogs for speech and grounding acts, which we performed using DAMSL (Dialog Act Mark-up in Several Layers).

The dialogs in this corpus are being used as a basis for the development of a new version of the TRIPS system that will be able to collaborate with users to solve disaster-handling tasks in the Monroe domain. In addition, the data is being used in research on the use of reference in language, and for research on content planning and natural language generation ([21]).

2 Related Work

2.1 Data collection

The domain and tasks used in the collection of this corpus are significantly more complex than those used in previous data collection efforts for task-oriented spoken dialog, such as the TRAINS corpus and the Maptask corpus ([13, 2]).

The Monroe domain is designed to encourage collaborative problem-solving and mixed-initiative interaction. The world in this domain, Monroe County, NY, has a large number of objects (e.g. roads, people, vehicles, crews and equipment). Emergencies include requests for medical assistance, car accidents, civil disorder, and larger problems such as flooding and snow storms. Resources at the user's disposal may include road crews, electric crews, ambulances, police units and helicopters.

Our tasks involved coordinating the solutions to several different problems, some of which interfered with each other and all of which were subject to time constraints. The tasks could not easily be solved in under ten minutes by one person working alone. Furthermore, we sometimes gave subjects new information when the dialog was almost finished that necessitated re-planning.

Care was taken to ensure that both participants had responsibility for the solution of the task, and to encourage mixed-initiative interactions. The range of tasks and complexity of the world allow for problem solving at different levels of granularity.

Video data is available that can be used to analyze how people use maps and make notes to aid planning.

2.2 Dialog annotation

In recent years there has been much research on annotation schemes for dialog. Traum and Hinkelman outline four levels of "conversational acts" in [22]. DAMSL, the scheme we used to annotate the Monroe corpus, annotates speech acts and some grounding acts. Previous work using DAMSL to annotate TRAINS-93 dialogs was reported in [9]; our work differs from that research because its focus is on content planning rather than on development of the annotation scheme. We have, however, benefited tremendously from that work. We have used the manuals and tools, and trained on previously-annotated TRAINS-93 dialogs.

Other recent research on annotation schemes for marking speech acts in dialog include the COCONUT project ([11, 10]), the HCRC Maptask annotation scheme ([4]), and the SWITCHBOARD-DAMSL project ([14]). Few of these projects have made available the kind of descriptive statistics given in this report, but where these facts are available we will compare them with our own results.

Previous researchers (e.g. [17], [16]) have analyzed texts to develop schemas for text generation, but did not apply these techniques to the analysis of dialog.

3 Dialog Collection

3.1 Subjects

There were twenty subjects (see table 1). All had completed at least three years of college. Fourteen were male; sixteen were native English speakers. Eight were undergraduates and eleven graduate students. One was not currently a student. Subjects were recruited from Computer Science and Linguistic undergraduates and Computer Science, Linguistics and Optics graduate students using newsgroups, personal invitation, and announcements made in classes. All subjects signed consent forms, and each was paid $7.50/hour for participating in the data collection.

Subject ID	Sex of subject	Native English speaker?	Graduate student (G)/ Non-graduate (NG)/ Other (O)	Task	Partner	Knew partner?
S1R	F	No	O	1	S1B	Yes
S1B	M	Yes	G	2	S1R	Yes
S2R	M	Yes	NG	4	S2B	No
S2B	M	Yes	G	3	S2R	No
S3R	F	No	NG	6	S3B	No
S3B	M	Yes	G	5	S3R	No
S4R	M	Yes	NG	8	S4B	Yes
S4B	F	No	NG	7	S4R	Yes
S5R	M	Yes	G	1	S4B	Yes
S5B	F	No	G	5	S5R	Yes
S6R	M	No	G	6	S6B	Yes
S6B	M	Yes	G	2	S6R	Yes
S7R	F	Yes	NG	7	S7B	Yes
S7B	M	Yes	NG	3	S7R	Yes
S8R	M	Yes	G	8	S8B	No
S8B	F	Yes	G	4	S8R	No
S9R	M	Yes	NG	8	S9B	Yes
S9B	M	Yes	NG	1	S9R	Yes
S10R	M	Yes	G	5	S10B	Yes
S10B	M	Yes	G	4	S10R	Yes

Table 1: Notes on subjects

Subjects worked in pairs. Some subjects knew each other while others did not.

3.2 Tasks and Domain

There were eight tasks, all disaster scenarios set in a (slightly idealized) Monroe County, New York state. The tasks were created by the experimenter and attempted by researchers in the University of Rochester before being assigned to subjects. They varied quite widely in difficulty, but even the simplest task was too complex to be understood and completed in ten minutes by a person working alone. (Interestingly, although these tasks are much more complex than those handled in many previous collection efforts, they still do not reflect the size of real-world disaster-handling tasks such as the snow storm that occurred in Rochester around the time of the data collection.)

The task descriptions themselves are given in appendix B.

3.3 Subject instructions

Each subject was given:

- A description of a task.
- Other helpful information (speeds and capacities of vehicles, locations of landmarks). This information was the same for all tasks.
- A map of Monroe county (shown in figure 2).

- Three pens of different colors.
- As much blank paper as needed.

Subjects were told to pretend that they worked at an emergency control center and had to construct a plan for their task. They were given 10 minutes to work on the plan on their own. Then, one subject in each pair (subject B) was instructed to set aside his/her working materials and was given a new map and some blank paper. The other subject (subject A) was told that he/she would be giving control of his/her task to B, who was at another site and knew nothing about the task. The subjects were to collaborate to complete the task, and then B was to summarize the solution for A. Full instructions are given in appendix D.

After the subjects had completed A's task, if there was time a new piece of information would be given to B and B would be asked to collaborate with A to incorporate this new information into the plan.

Then, A would be asked to set aside his/her working materials and both subjects would collaborate to solve B's task. The same instructions as were given at the beginning would be repeated or summarized before starting the second dialog. After completion of the second dialog, subjects were given the opportunity to ask questions or make comments.

3.4 Experimental setup

The data collection was held in the speech lab at the University of Rochester. Subjects sat at two tables with a barrier between them so that they could not see each other. We set up the barrier for several reasons. We wanted to discourage use of body language because of the impossibility of recording it completely. We also wanted to encourage a variety of types of language use; that is, to discourage subjects from just sharing maps and pointing.

Subject R in each pair was recorded with a Sennheiser HMD 414 microphone and subject B with a Sennheiser HMD 224 microphone. In general, each subject could hear his/her own voice and that of the partner in the headphones. Sometimes, however, the connection was accidentally broken for a short period of time and where this occurred it generally shows up in the transcript as a comment by the subject. The audio data was recorded onto DAT tapes at 48 kHz. The subjects' hands (their gestures and their use of the maps) were videotaped.

An experimenter was always in the room with the subjects. The experimenter gave instructions, set recording levels, and ran the video recorders. The experimenter could also listen to the subjects with headphones plugged into the DAT machine.

3.5 Comments on the Dialogs

All dialogs ended with the participants thinking that they had solved every part of the task. While this was generally true, in one case, participants failed to solve the extra problem given at the end, and in another case, participants invented extra police units and ambulances to help them solve the task (see appendix A).

Subjects did not limit themselves to discussing the task. They made meta-comments about the domain or task, occasionally commented if they couldn't hear the other participant, and sometimes made jokes to each other. In addition, because many of them are familiar with Monroe county, some subjects made use of world knowledge not given in the task description or shown on the map.

In order to test our hypothesis that a lack of grounding does significantly affect dialog, we had two of our subjects return and conduct two dialogs each with one experimenter, solving tasks they had not previously solved. In the first dialog, the experimenter behaved as a human would, processing

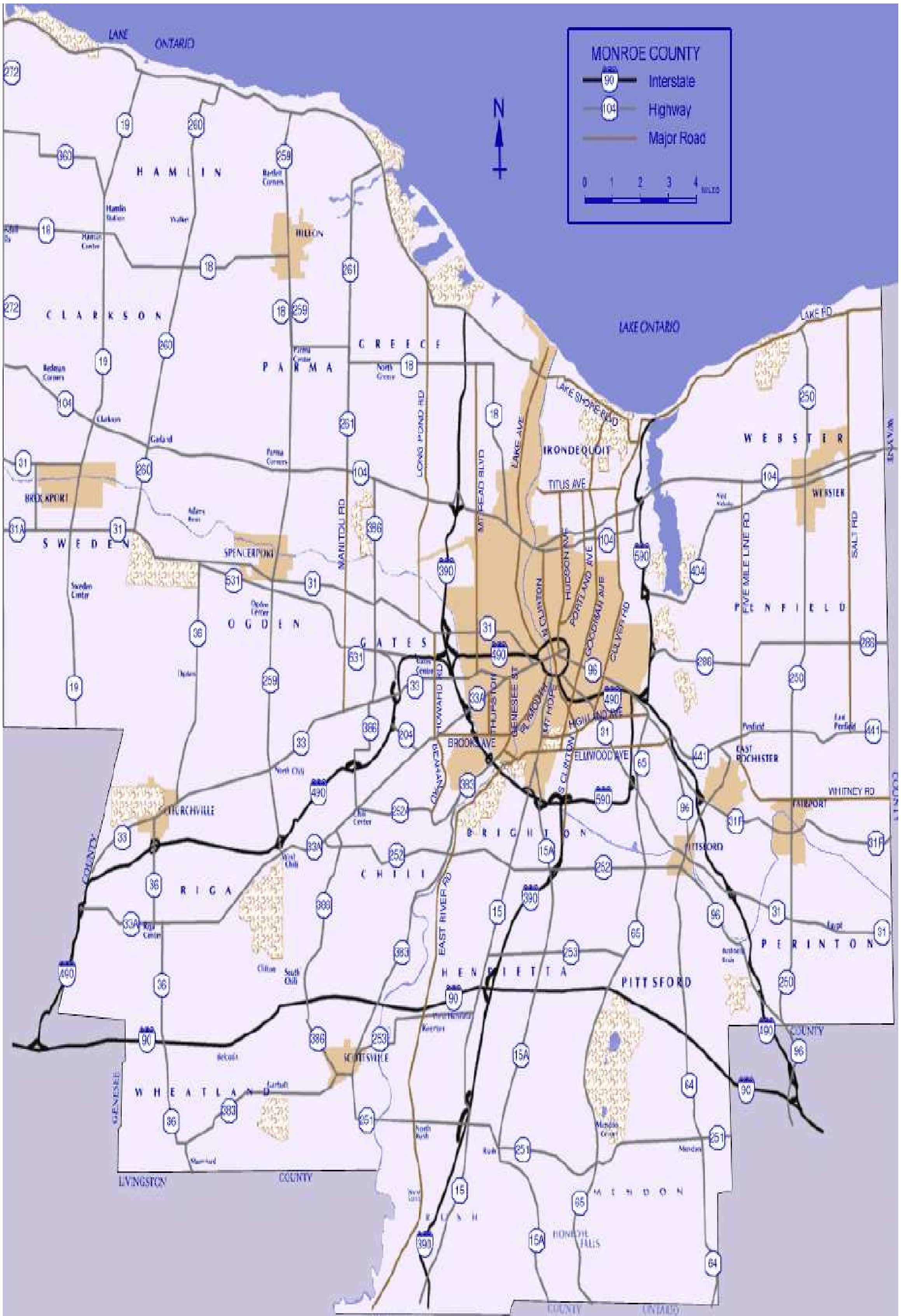

Figure 2: Map of Monroe county used in dialog collection

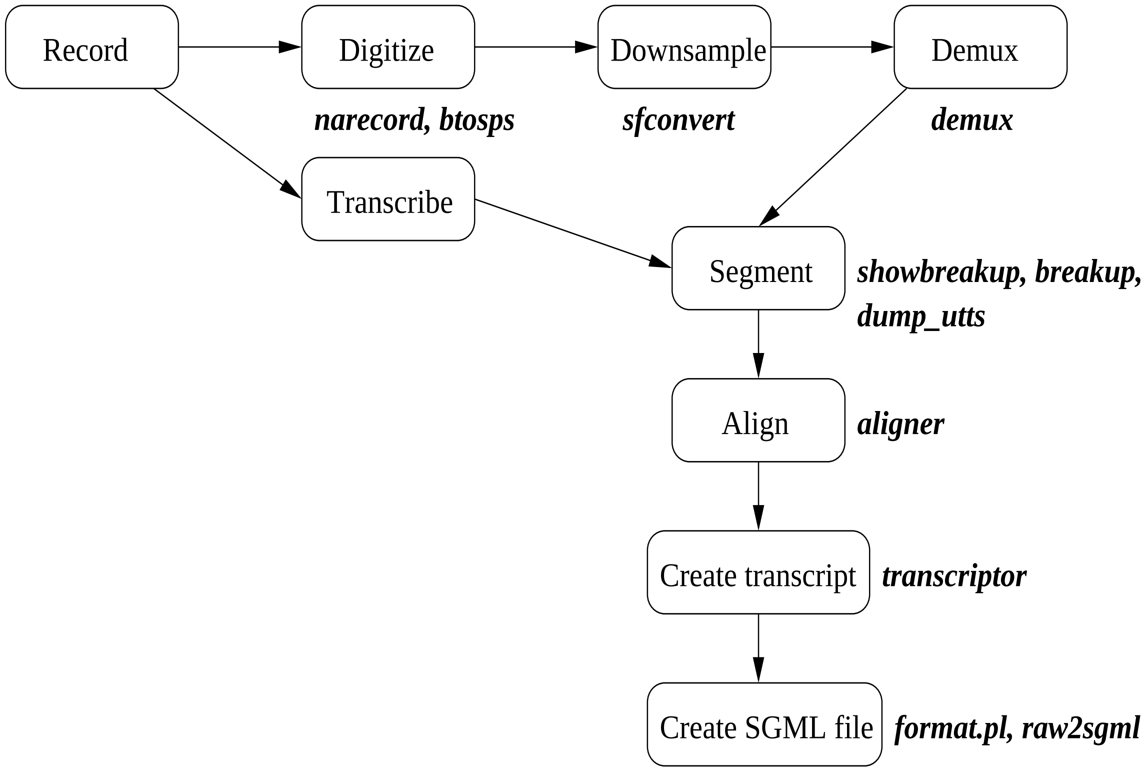

Figure 3: Initial processing of Monroe dialogs

speech in real-time and providing feedback. In the second, the experimenter did not back-channel, perform grounding acts or interrupt. The people took longer to solve the task in the second dialog, paused longer between turns, and said they thought the experimenter was not paying attention or not understanding. One in particular exhibited a great deal of frustration.

These dialogs do not form part of the Monroe corpus. They provide merely anecdotal evidence supporting our claims in the context of human-human dialog; our ideas have to be tested in the context of human-computer dialog to determine whether it is indeed substantially the same as human-human dialog. We have no intentions at this time of collecting further dialogs of this sort, although we think it would be very interesting to do so.

4 Dialog Segmentation, Transcription and Alignment

The process from raw audio to annotated SGML files is outlined in figure 3. Here we go into more detail.

First, each dialog was recorded from the DAT machine at 44.1kHz using the command **narecord -s 44100 <filename>.raw**. The resulting audio file had no header, so it was converted to a file with an ESPS header using **btosps -f 44100 -n 2 <filename>.raw <filename>.fea**, and then any blank space at the beginning and end of the files was removed. To enable us to use many of the ESPS tools (in particular, the aligner) we downsampled the audio to 16kHz using **sfconvert -s16000 <filename>.fea dialog.fea**. The resulting 2-channel file was separated into two single channel files using **demux -e0,1 dialog.fea speaker0.fea speaker1.fea** so that we could do the segmentation.

We used **showbreakup** to manually segment the audio into utterances. Each dialog was segmented by one annotator using the guidelines in [12]; the segmentations were checked during the early stages of annotation. The segmentation breaks the audio data up not only into utterances, but also into turns separated by speaker changes. Not every speaker change results in a new turn; for instance, back-channels do not cause a turn break. Following the segmentation, **breakup** was used to break the audio file into individual utterance files.

We segmented the dialogs into utterances for the following reasons:

- To simplify the alignment process.
- To enable us to more easily use the DAMSL annotation tool DAT.

No segmentation is perfect for every type of processing or dialog annotation. It is frequently necessary to re-segment to some extent for different analyses. Furthermore, a lot of information is left out if one only looks at the utterances of a segmented dialog; pauses, overlap, breaths, and non-verbal sounds may all be missing or slightly incorrect in the transcription. Listening to the entire dialog at least once can therefore be very helpful for annotators.

Each dialog was transcribed by hand from the audio. We used the script **trtoutts** to break up the text transcription into individual utterance text files matching the audio segmentation. Then, we used the ESPS aligner to align all the individual utterance files. This produced .words and .phones files for each utterance. The alignments were later checked and corrected by hand.

After the alignment process had finished, we used **dump_utts -e words '*' speaker0 speaker1** to create a .words file for each speaker. Then we used **transcriptor** to get a text transcript including silences from the .words files, and **format.pl** and **raw2sgml** to get the .sgml transcript file.

The **narecord** program is part of the DAT machine software. The tools we used, including **btosps, sfconvert, demux** and the aligner, are from the ESPS/*waves*+ 5.3 package. The scripts we used were derived from those explained in detail in [13], but were not identical. We had to update or modify them for use with our longer dialogs and prevent the deletion of files we wanted to keep.

We have not annotated or modified the video data in any way. We know of no annotation schemes for video data, and have not chosen to focus on it in our research.

Errata should be sent to *stent@cs.rochester.edu*.

5 Dialog Annotation using DAMSL

Our initial annotation of the Monroe corpus used DAMSL to annotate for speech and grounding acts. Other annotations will be reported in future papers.

5.1 DAMSL annotation scheme

DAMSL is a system for annotating aspects of utterances in dialog such as illocutionary acts, topic, and relationships between utterances. It came out of the Discourse Research Initiative meetings in 1996, 1997 and 1998; the manual we used dates from 1997 [1][2]. The set of acts outlined in DAMSL are intended to provide a way for different research groups to annotate data and then share it [9].

[2] In 1998 DAMSL was renamed BF; no manual for the updated annotation scheme is currently publicly available.

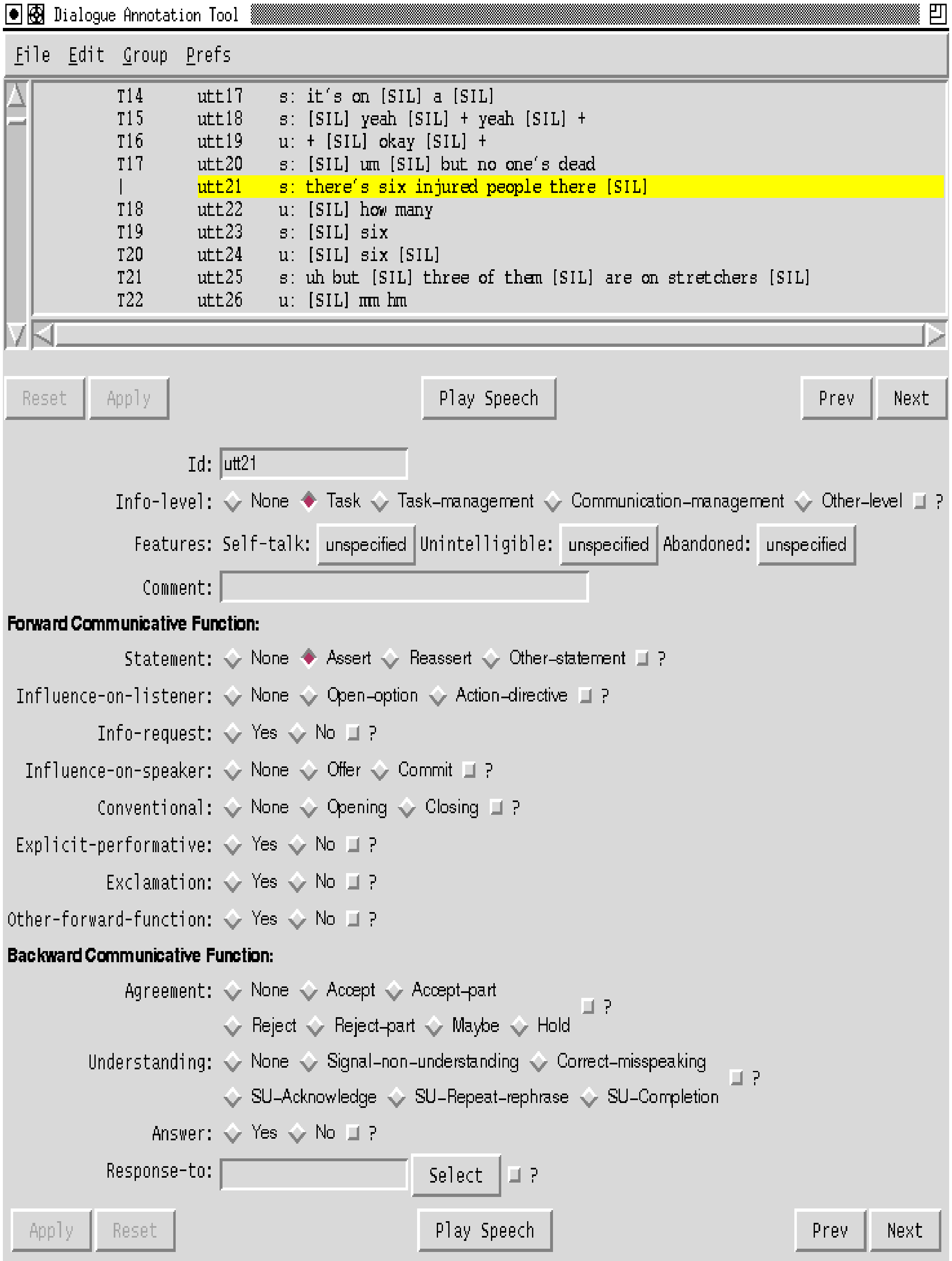

Figure 4: Screenshot of DAT tool showing DAMSL tags

Dimension	PA	PE	kappa	sig
Influence-on-listener	0.973	0.770	0.883	< 0.01
Influence-on-speaker	0.954	0.730	0.831	< 0.01
Info-request	0.981	0.811	0.902	< 0.01
abandoned	0.939	0.820	0.663	< 0.01
agreement	0.955	0.601	0.888	< 0.01
answer	0.979	0.854	0.856	< 0.01
conventional	0.999	0.988	0.882	< 0.01
exclamation	0.994	0.981	0.699	< 0.01
info-level	0.877	0.702	0.588	< 0.01
other-forward-function	0.988	0.901	0.881	< 0.01
performative	0.996	0.992	0.452	0.025
response-to	0.899	0.397	0.832	< 0.01
statement	0.926	0.409	0.876	< 0.01
understanding	0.962	0.559	0.914	< 0.01
unintelligible	0.983	0.976	0.308	< 0.01

Table 2: Kappa scores for annotation of eight Monroe county dialogs. N = 2897; k = 2

DAMSL tags fall into 13 dimensions that are in three layers: forward looking functions, backward looking functions, and information level. All three layers plus miscellaneous tags are shown in figure 4. Tag dimensions that involve forward-looking functions indicate the type of illocutionary act the utterance is performing, e.g. assertions, info-requests, action-directives or commits. Tag dimensions that involve backward-looking functions indicate how the utterance relates to previous utterances and include answers to questions, indications of degree of understanding, and indications of (dis-)agreement. Information-level tags indicate the dialog level of the contents of the utterance (task, task-management or communication-management). Each tag dimension is marked separately from the others. We will describe the individual dimensions in section 6 below.

DAT and the DAMSL manual are available at the DAMSL web page at:
http://www.cs.rochester.edu/research/trains/annotation/

Process of annotation

Eight of the twenty dialogs in the Monroe corpus were selected for annotation: s2, s4, s6, s12, s14, s16, s17, and s18. They vary in length from 163 to 647 utterances (7 to 27 minutes) and include a variety of tasks and speakers.

There were two annotators. Each had limited experience annotating TRAINS-93 dialogs and checking her annotations against a reference annotation. Each also had the DAMSL manual [9].

First, the two annotators agreed on the segmentation. Each annotator then annotated all eight dialogs using the DAT tool [8]. The annotators did not discuss aspects of the annotation until after annotating all eight dialogs. Then, one annotator made an initial reconciled annotation, and the two met and discussed differences to arrive at consensus.

Kappa scores for the annotation are given in table 2. These scores measure inter-annotator agreement for each tag dimension across all eight dialogs using the method from [7]. With a kappa score above .67 one's results are tentative; with one above .8 they are considered to be reliable [3]. Most of these kappa scores are reliable or tentatively reliable. The *info-level* dimension has a low kappa score because the annotators disagreed about which utterances were task-management utterances. The *performative* and *unintelligible* dimensions have low kappa scores because there was very little data for these dimensions.

5.2 Comments on DAMSL

The two annotators found the decision trees and examples in the DAMSL manual very helpful for clarifying difficult utterances. However, there were some less-than-intuitive decisions in these trees. For instance, one might ask why an accept-part is always preferred over a reject-part.

There is a distinct possibility of annotators confusing surface form with speaker intent, for instance labeling an info-request in the form of a statement as an other-statement. These types of confusions show up in the results reported here. We are not sure that this was the intent of the designers of the scheme.

In this domain, some of the DAMSL dimensions do not seem truly independent. The *Open-option* and *Offer* tags almost always co-occurred, and the *SU-Acknowledge* and Accept tags frequently co-occurred. In this type of domain, it might be possible or desirable to combine some tag dimensions, and DAMSL is supposed to be capable of modification to specific domains or purposes, as in SWITCHBOARD-DAMSL ([14]), or in current work at the National University of Mexico extending DAMSL to cover gesture as communication ([18]).

Finally, the annotators found it difficult to distinguish between the task and task-management levels of dialog. This confusion is reflected in the lack of agreement between annotators; what is not reflected is that the annotators themselves were not sure of their own annotations in this area at times.

6 Characteristics of the Monroe corpus

As we stated previously, we are trying to determine the relative frequency of dialog acts, how they relate to each other, and how they can be combined to form dialog contributions. The initial questions we wanted to answer are:

1. How often do various acts occur?
2. Which combinations of acts are often realized in a single utterance?
3. Which acts are produced early in a turn, and which later?
4. What sequences of acts most commonly occur?

We use the annotated dialogs to answer these questions. First, we consider individual acts (e.g. *Assert, Info-request*), examining how often they occur, where in a turn they occur, and with what other acts they most frequently occur. Then, in order to be able to examine sequences of combinations of acts, we give some simple bigram models obtained from our data. We can use these kinds of information to

These answers can inform the development of our theory of content planning.

There are drawbacks to purely descriptive analyses. The information contained in the annotation must be interpreted and then hand-coded into any implementation of the theory. Also, it is difficult to reason about the combinations of intentions and ordering of acts simultaneously. For these reasons, we are also producing probabilistic dialog models suitable for use in prediction and planning of dialog contributions.

The type of dialog affects the answers to questions like these; for instance, in tutoring dialogs the number of questions will be much higher than in general task-related dialogs. These results should only be taken to apply to task-oriented dialogs, and the relative sparseness of data (2900 utterances is not a large corpus) should also be taken into account.

Tag	Number of occurrences			
	Single-utt. turns	Multiple-utt. turns, first utt.	Multiple-utt. turns, last utt.	Total (percent)
Assert	385 (154)	215 (103)	323 (223)	1149 (39.6%)
Reassert	73 (51)	51 (38)	88 (70)	290 (10.0%)
Other-statement	10 (9)	8 (7)	13 (8)	38 (1.3%)
Action-directive	73 (3)	38 (0)	64 (1)	230 (7.9%)
Open-option	39 (0)	25 (1)	46 (0)	140 (4.8%)
Info-request	116 (83)	28 (23)	116 (87)	283 (9.8%)
Commit	130 (0)	67 (1)	75 (0)	341 (11.8%)
Offer	41 (1)	24 (0)	46 (0)	141 (4.9%)
Other-forward-function	53 (49)	51 (46)	25 (22)	165 (5.7%)
Explicit-performative	2 (1)	2 (0)	3 (2)	9 (0.3%)
Exclamation	16 (4)	6 (3)	3 (2)	34 (1.2%)
Opening	2 (2)	7 (7)	0 (0)	9 (0.3%)
Closing	6 (2)	0 (0)	3 (2)	9 (0.3%)
Accept	339 (0)	198 (1)	65 (0)	661 (22.8%)
Accept-part	1 (0)	3 (0)	1 (0)	5 (0.2%)
Reject	6 (0)	6 (0)	0 (0)	15 (0.5%)
Reject-part	3 (0)	2 (0)	1 (0)	7 (0.2%)
Maybe	4 (0)	1 (0)	1 (0)	7 (0.2%)
Hold	23 (0)	10 (0)	0 (0)	34 (1.2%)
Signal-non-understanding	7 (0)	4 (0)	2 (0)	14 (0.5%)
Correct-misspeaking	6 (0)	0 (0)	1 (0)	10 (0.3%)
SU-Acknowledge	367 (50)	228 (42)	57 (3)	702 (24.2%)
SU-Repeat-rephrase	45 (4)	32 (9)	10 (0)	97 (3.3%)
SU-Completion	9 (2)	7 (1)	2 (0)	21 (0.7%)
Answer	138 (2)	63 (1)	26 (0)	242 (8.3%)

Table 3: Frequency of occurrence of the most frequently-occurring tags, with position

6.1 Frequency of tags with position in turn, and combinations of tags

The eight dialogs contain a total of 1758 turns, 2899 utterances. There are 19328 words not including non-speech markings. There are 1069 unique words. There are 2873 utterances that are not directed at the experimenter; we include only these utterances in discussion from now on.

There are 26 turns (40 utterances) that contain no speech, e.g. isolated coughs, clearing of throat. 16 utterances are certainly unintelligible, and another 12 might be.

Twenty-two utterances are certainly self-talk, and another 25 might be. This is a relatively small amount of self-talk; possibly subjects talked more to themselves in the ten minutes when they were working on their own.

Two hundred twenty-one utterances are marked as abandoned; another 19 as possibly abandoned. This is a high number of abandoned utterances, and probably an over-estimate. The DAMSL manual states that an utterance should be marked as abandoned only if it serves no purpose in the dialog, not if it is simply incomplete. However, the annotators marked some incomplete utterances as abandoned.

There are 1075 single-utterance turns and 656 multiple-utterance turns (406 two-utterance turns). Table 3 gives the frequencies of the different acts, with their relative positions within turns. The first number in each pair is the total count of relevant utterances with that tag; the second is the count of those utterances having no other forward- or backward-looking function tags.

An utterance may satisfy more than one intention; DAMSL allows us to mark each tag dimension independently so we can account for several possible intentions behind each utterance. We will now look at those tags that occurred at least ten times in the dialogs, with their most frequently-occurring combinations.

Statements

Statements are utterances that make a claim about the world. The test for a statement, in general, is whether it can be responded to with a "That's true" or "That's not true". If the speaker is trying to change the beliefs of the hearer, then the statement is an *Assert* or *Reassert*; otherwise, it is an *Other-statement*.

Assert Simple assertions are the most commonly occurring acts in these dialogs. They occur at all points in turns, but a little less frequently as the initial utterance in a multiple-utterance turn. The contents of and relationships between these assertions that have no other tags are things that it would be very helpful to know in order to design a generation component; we will address these questions in an analysis of rhetorical structure in future work.

The *Assert* tag most frequently co-occurs with the *Answer* tag at the beginning of a turn; we will see later that answers are most frequently realized as assertions. Likewise, assertions are frequently used as a vehicle for producing action-directives and open-options anywhere in a turn. At the end of a turn, they may be used to produce info-requests, but this is not common. They are rarely used in combination with backward-looking functions. *Assert* is more frequently combined with other tags at the end of a multiple-utterance turn than at the beginning.

Type of utterance	Most frequently occurring combinations with Assert
single-utterance turn	Answer (106) Action-directive, Commit (40) Open-option, Offer (25) Exclamation (7) Commit, Action-directive, Answer (4)
multiple-utterance turn, first utterance	Answer (47) Action-directive, Commit (20) Open-option, Offer (17) Accept, SU-Acknowledge (4)
multiple-utterance turn, last utterance	Open-option, Offer (31) Action-directive, Commit (29) Answer (18) Info-request (7)

Table 4: Tags that co-occur with *Assert*

Reassert Like *Assert*, the *Reassert* tag occurs at all points in a turn. It is more frequently combined with other tags than the *Assert* tag, but there are far fewer *Reasserts* than *Asserts*. As we would expect, it is not as frequently combined with the *Answer* tag as is *Assert*. It is sometimes combined with the *Info-request* tag; these are confirmation questions:

Example 1

A [SIL] So [SIL] you said [SIL] two fifty two [SIL]

Type of utterance	Most frequently occurring combinations with Reassert
single-utterance turn	Answer (7) Action-directive, Commit (5) Info-request (4)
multiple-utterance turn, first utterance	Action-directive, Commit (5) Answer (3)
multiple-utterance turn, last utterance	Action-directive, Commit (7) Info-request (5) Answer (3) Open-option, Offer (2)

Table 5: Tags that co-occur with *Reassert*

Other-statement Utterances marked with this tag fall into four broad groups: comments on the planning process (e.g. “That’s good”, “And just to add spice to this”, “So that doesn’t help us much”), utterances that may be self-talk or that are borderline unintelligible (muttering, out-of-context utterances) and remarks about the speaker’s behavior or mental state (e.g. “I haven’t been keeping track of that”, “I guess so”, “I thought it could only pull trees”). All of these take the form of statements, but their truth either can’t be determined or is unknown at the time of utterance. This tag was infrequently used and was used in combination with a variety of other tags; the only emerging pattern is the case where questions phrased as statements are marked *Other statement*, e.g.:

Example 2

A [SIL] Okay [SIL] where where [SIL] that’d be [SIL]
where Genesee Brooks and [SIL] three eighty three + connect +

Type of utterance	Most frequently occurring combinations with Other-statement
single-utterance turn	none
multiple-utterance turn, first utterance	none
multiple-utterance turn, last utterance	Info-request (4)

Table 6: Tags that co-occur with *Other-statement*

Type of utterance	Most frequently occurring combinations with Action-directive
single-utterance turn	Assert, Commit (40) Commit (16) Reassert, Commit (5)
multiple-utterance turn, first utterance	Assert, Commit (20) Commit (10) Reassert, Commit (5)
multiple-utterance turn, last utterance	Assert, Commit (29) Commit (20) Reassert, Commit (7)

Table 7: Tags that co-occur with *Action-directive*

Influence-on-listener and Influence-on-speaker

The influence-on-listener dimension includes the tags *Action-directive* and *Open-option*. An utterance would be labeled with one of these tags if it created an obligation for the hearer to do something, that is, if it could be responded to with an "I can't do that". These two tags differ in the strength or force of the obligation created; *Action-directive* is used for commands, *Open-option* for suggestions. If the hearer can ignore the utterance, then it is probably not an *Action-directive*.

The influence-on-speaker dimension includes the tags *Offer* and *Commit*. Both express an intention by the speaker to perform some action. The difference between an *Offer* and a *Commit* is that the first is conditional on the hearer's acceptance, and the second is not.

Action-directive There is a large number of influence-on-speaker and influence-on-listener tags in this corpus, because it is a planning domain. Because the planning is collaborative, most of the time these two dimensions are both marked where one of them is marked.

Action-directives occur in single-utterance turns, at the end of multiple-utterance turns, and occasionally as sequences of commands in multiple-utterance turns. Examples of *Action-directives* include:

Example 3
A Take it over to Clinton
+ [SIL] And then come back up Mount Hope +
Example 4
A So [SIL] um let's send the digging [SIL] truck [SIL]
to [SIL] the elmwood bridge [SIL]
A [SIL] Uh [SIL] and get them started on that

Many *Action-directives* are also marked *Assert* and *Commit* (a proposal for a shared action is taking place). Those that are marked *Assert* are phrased as declarations of future action, e.g. "So we have to wait for the ambulance to come back". Those that are marked only *Action-directive, Commit* are phrased as commands, e.g. "At the same time send both ambulances to the airport". Those that are not marked *Commit* are typically instructions for the hearer to find some place on the map or perform some communicative action, e.g. "Go all the way over to East River road", "You wanna read back your plan just to make sure we've got this straight".

Type of utterance	Most frequently occurring combinations with Commit
single-utterance turn	SU-Acknowledge, Accept (52) Assert, Action-directive (40) Action-directive (16) Reassert, Action-directive (5)
multiple-utterance turn, first utterance	Accept, SU-Acknowledge (23) Assert, Action-directive (20) Action-directive (10) Reassert, Action-directive (5)
multiple-utterance turn, last utterance	Assert, Action-directive (29) Action-directive (20) Accept, SU-Acknowledge (7) Reassert, Action-directive (7)

Table 8: Tags that co-occur with *Commit*

Commit *Commits* occur most frequently in single-utterance turns; they tend to immediately follow *Action-directives* or *Open-options*.

The reason *Commit* more frequently occurs with *SU-Acknowledge, Accept* than with *Assert, Action-directive* is because commitments to action-directives and to open-options are both marked *Commit*, and because people sometimes agree to some action more than once (possibly for emphasis or clarity).

Open-option and Offer

There were fewer *Open-options* in this data than there were *Action-directives*, but like the *Action-directive* tag, the *Open-option* tag occurs in single-utterance turns or at the end of multiple-utterance turns. Examples include:

Example 5

A You think it makes sense uh in the interim
there to dispatch at least one electrical [SIL] crew [SIL]
over to help the first one

Example 6

A Meanwhile [SIL] we can send one of the electrical crews
[SIL] to three [SIL] eighty three and brooks

The *Offer* tag always co-occurs with *Open-option*, and vice versa. In this domain, almost all proposals are proposals for shared action. There were a few *Action-directives* that were not proposals for shared action ; by contrast, there were no *Open-options* that were not proposals for shared action. We expect that in a tutoring or explanation-generation domain, there would be more *Open-options* that were not also *Offers*, e.g. "You could try factoring the equation before solving it", or "You could use vinegar and warm water to take wallpaper off the walls because it dissolves the adhesive".

Type of utterance	Most frequently occurring combinations with Open-option
single-utterance turn	Assert, Offer (25) Offer (9)
multiple-utterance turn, first utterance	Assert, Offer (17) Offer (5)
multiple-utterance turn, last utterance	Assert, Offer (31) Offer (7) Info-request, Offer (6)

Table 9: Tags that co-occur with *Open-option*

Type of utterance	Most frequently occurring combinations with Offer
single-utterance turn	Assert, Open-option (25) Open-option (9)
multiple-utterance turn, first utterance	Assert, Open-option (17) Open-option (5)
multiple-utterance turn, last utterance	Assert, Open-option (31) Open-option (7) Info-request, Open-option (6)

Table 10: Tags that co-occur with *Offer*

Info-request

An *Info-request* is an utterance that produces an obligation to answer. This dimension includes all questions. *Info-requests* do not have to appear syntactically as questions; if the intonation is that of a question but the syntax that of a statement the utterance is an *Info-request*, and if the hearer answers that is a clear indication of an *Info-request* regardless of syntactic structure.

There are 293 *Info-requests* in the eight dialogs (table 3). Most of these occur right before a speaker change; those that don't are sequences of questions, rhetorical questions, abandoned questions or questions that the speaker asks him/herself.

Most *Info-requests* in these dialogs are simple questions, although some are indications of lack of understanding or requests for clarification and some are indirect commands. It seems that a request

Type of utterance	Most frequently occurring combinations with Info-request
single-utterance turn	SU-Repeat-rephrase, Hold (6) Reassert (4) Signal-non-understanding (3)
multiple-utterance turn, first utterance	none
multiple-utterance turn, last utterance	Assert (7) Open-option, Offer (6) Reassert (5)

Table 11: Tags that co-occur with *Info-request*

Type of utterance	Most frequently occurring combinations with Other-forward-function
single-utterance turn	none
multiple-utterance turn, first utterance	SU-Acknowledge (3)
multiple-utterance turn, last utterance	Info-request (2)

Table 12: Tags that co-occur with *Other-forward-function*

Type of utterance	Most frequently occurring combinations with Exclamation
single-utterance turn	Assert (7)
multiple-utterance turn, first utterance	Assert (2)
multiple-utterance turn, last utterance	none

Table 13: Tags that co-occur with *Exclamation*

for information imposes such a burden to respond on the listener that, in general, no other act should be combined with this one.

Other-forward-function

The *Other-forward-function* tag marks mostly turn-management utterances such as prompts (e.g. "Right?"), turn-taking utterances (e.g. certain uses of "Okay"), and turn-keeping utterances like "Just a sec". It is intended to also capture other forward-looking functions not handled in any other dimension, but we do not think we found any of these.

The *Other-forward-function* tag occurred mostly on its own.

Exclamation

The *Exclamation* tag is self-explanatory.

There were very few exclamations in this corpus, most occurring in single-utterance turns. Fourteen came from one dialog, s2. Examples include "Bummer, man", "You need to land them somewhere!", "Oh no I'm sorry" and "Excellent!".

If the *Exclamation* tag occurred with another tag, that tag was most frequently *Assert*.

Agreement

When a speaker has just heard an offer, a command, or a statement, there is usually a kind of obligation to respond with either agreement or disagreement. These responses are marked using this dimension. The agreement tags include *Accept, Accept-part, Reject, Reject-part, Maybe* and *Hold*. This annotation scheme provides no means of marking implicit (dis-)agreement, because only actual utterances can be marked.

Accept The *Accept* tag is the third most frequently-marked tag in this corpus. It occurs mostly in the initial utterance of a turn. Frequently, it occurs with the *SU-Acknowledge* tag, but sometimes with the *SU-Repeat-rephrase* tag. Also, sometimes an *Accept* can be used to commit to a course of action.

Type of utterance	Most frequently occurring combinations with Accept
single-utterance turn	SU-Acknowledge (228) SU-Acknowledge, Commit (52) SU-Repeat-rephrase (24)
multiple-utterance turn, first utterance	SU-Acknowledge (139) Commit, SU-Acknowledge (23) SU-Repeat-rephrase (14)
multiple-utterance turn, last utterance	SU-Acknowledge (39) Commit, SU-Acknowledge (7) SU-Repeat-rephrase (5)

Table 14: Tags that co-occur with *Accept*

Examples of *Accepts* include:

Example 7
A Well let me just like [SIL] double check to make sure we have ev all the people
B [SIL] Okay [SIL]

Example 8
A [click] [SIL] We should [SIL] have the vans [SIL] pick up the people at two sixty one [SIL] and three ninety cause there's nine each [SIL]
B Right

Example 9
A So I'll cross off bus C
B Okay

Those places where the *Accept* tag does not occur with any tag in the *Understanding* dimension are probably annotator errors.

Hold The *Hold* tag indicates that the speaker does not want to commit to agreeing or disagreeing to a prior utterance; usually there is some communication problem such as an unclear referent. In human-computer dialogs, because of speech recognition errors there tend to be quite a few *Hold* utterances.

All *Hold* tags occurred at the start of turns. This tag occurred with a variety of tags in the **Understanding** dimension, and also with the *Info-request* tag. Examples include:

Example 10
A Because you got to fix the tree first
B Mmm

Example 11
A [SIL] One [SIL] thirty can go pull a tree or + fix a line +
B + One thirty + what do you mean [SIL] one thirty [SIL] the the [SIL] + time +

There were not very many occurrences of the *Hold* tag, so these results are tentative.

Type of utterance	Most frequently occurring combinations with Hold
single-utterance turn	SU-Acknowledge (7) Info-request, SU-Repeat-rephrase (6) Info-request, SU-Acknowledge (2) Info-request, Signal-non-understanding (2)
multiple-utterance turn, first utterance	SU-Acknowledge (6) SU-Repeat-rephrase (2)
multiple-utterance turn, last utterance	none

Table 15: Tags that co-occur with *Hold*

Type of utterance	Most frequently occurring combinations with SU-Acknowledge
single-utterance turn	Accept (228) Accept, Commit (52)
multiple-utterance turn, first utterance	Accept (139) Accept, Commit (23)
multiple-utterance turn, last utterance	Accept (39) Accept, Commit (7)

Table 16: Tags that co-occur with *SU-Acknowledge*

Understanding

The understanding dimension includes the *Signal-non-understanding, Correct-misspeaking, SU-Acknowledge, SU-Repeat-rephrase* and *SU-Completion* tags. It is used to mark utterances that indicate the level of understanding the speaker has of an utterance he/she has recently heard. This annotation scheme does not give a way of annotating implicit acknowledgment; only explicit signals of understanding can be marked. The understanding (and agreement) dimensions therefore do not give a complete model of grounding behavior. but do provide a good start.

The *Signal-non-understanding* tag indicates a lack of understanding. The other tags all indicate some degree of understanding, with *SU-Acknowledge* being weaker than the other forms.

SU-Acknowledge There are 709 *SU-Acknowledge* acts. The vast majority of these occur right after a speaker change; in fact, they account for 1/3 of initial utterances in turns. The very few that occur later in a turn are in multiple-utterance turns that start with or consist of sequences of acknowledgments, or that start with something that is non-linguistic, such as laughter or a sigh. These results indicate that we should put acknowledgments at the start of a turn, or in a turn consisting only of the acknowledgment. Also, the sheer number of acknowledgments is an argument that a conversational agent should be able to plan to produce these in a principled way.

The *SU-Acknowledge* tag generally co-occurs with the *Accept* tag. Most utterances labeled *SU-Acknowledge* have no forward-looking function; this may be due to the principle of negative evidence, which states that an utterance is to be assumed to be in common ground unless there is evidence to the contrary [6]. The large number of *Commits* co-occurring with *SU-Acknowledgments* is an artifact of the cooperative-planning nature of these dialogs.

Type of utterance	Most frequently occurring combinations with SU-Repeat-rephrases
single-utterance turn	Accept (24) Hold, Info-request (6) Accept, Commit (3)
multiple-utterance turn, first utterance	Accept (14) Hold (2) Accept, Commit (2)
multiple-utterance turn, last utterance	Accept (5) Accept, Commit (2)

Table 17: Tags that co-occur with *SU-Repeat-rephrase*

SU-Repeat-rephrase Most occurrences of the *SU-Repeat-rephrase* tag were at the beginning of turns. They were frequently combined with other tags, especially the *Accept* tag, e.g.:

Example 12

A [SIL] And bring them to Saint Mary's
B To Saint Mary

This tag also occurred with the *Commit* tag, as in:

Example 13

A Go to Genesee street and uh [SIL] cut up the tree
B [SIL] [lipsmack] [SIL] Oh
B Mmm cut tree

Finally, this tag occurred with the *Info-request* tag, e.g.:

Example 14

A Yeah we [SIL] we are sending them with the [SIL] truck that can be used to fix [SIL] both [SIL] lines and the [SIL] tree
B Oh the one that can do both?

Sometimes, utterances marked with this tag are also marked as possibly being self-talk.

SU-Repeat-rephrase therefore serves at least two very different purposes: to clarify something, and to agree with something strongly. The form and intonation are what distinguish one purpose from another; a generation component must plan carefully so as not to produce very ambiguous utterances of this type.

Answer

An answer is a response to an *Info-request*. Utterances that postpone answering, for instance clarification questions, are not marked *Answer*.

There was a fairly large number of answers in this corpus. The majority occurred at the start of a turn; those that didn't were usually preceded by an utterance to take or keep the turn, e.g. "Just a minute"; or by a restatement of the facts in the question.

Most answers were some sort of assertion. Some, however, were also action-directives:

Type of utterance	Most frequently occurring combinations with Answer
single-utterance turn	Assert (106) Reassert (7) Accept, Assert, SU-Acknowledge (4) Commit, Action-directive, Assert (4) Accept, Assert (3) Assert, Reject (2)
multiple-utterance turn, first utterance	Assert (47) Assert, Accept, SU-Acknowledge (4) Reassert (3) Reject, Assert (2)
multiple-utterance turn, last utterance	Assert (18) Reassert (3) Commit, Action-directive, Assert (2)

Table 18: Tags that co-occur with *Answer*

Example 15

A	And then where is it going to go
B	Um [SIL] then it should go to Culver road

Other

The tags *Explicit-performative, Opening, Closing, Reject, Accept-part, Reject-part, Maybe, Signal-non-understanding, SU-Completion*, and *Correct-misspeaking* occurred so rarely in this corpus that we can make no inferences about how they might be used.

6.2 Sequences of (combinations of) tags

Producing statistics about sequences of acts from DAMSL annotations can be difficult, because one has to select one act from the set of acts labeled for each utterance. A finite-state machine, with the arcs labeled with combinations of tags, captures both the multi-dimensionality of the annotation scheme and the sequencing of the utterances and turns.

We have used our annotations to produce finite state machines that model some aspects of the dialog behavior in our corpus. In these machines, each arc from state S_{i-1} to state S_i corresponds to a possible combination of tags for one utterance T_i; each arc is also weighted with the transition likelihood, that is, the fraction $\frac{number of times S_{i-1} produced T_i}{number of times S_{i-1} produced any T}$[3].

We have produced several different models, among them some that include the information level and others that don't; some that include turn information and others that don't; and some that include only utterance sequences that were seen more than a certain number of times.

These models could be made much more useful with the addition of information about argumentation acts and utterance content. They suffer from the same incompleteness problems (resulting from sparseness of data) that plague language models. However, they provide a nice summary of the complex interactions involving the different dimensions that can be annotated and the sequencing

[3]Essentially, these models are bigrams. The use of more context would be good, but due to data sparseness and lack of an annotation for dialog structure, we have not yet explored this.

information that gives us dialog coherence. We can use them to identify patterns that appear in dialog, and to build interpretation and content-planning modules.

The FSM constructed from all the utterances in all eight dialogs is very large. For that reason, we have chosen to look at several parts of it that relate to the more frequently-occurring tags. For each of the FSMs below, we have extracted from the overall FSM the arcs, states and transition probabilities that include the utterances in the turns immediately before or after turns containing the relevant tag. Turn information is included. The information-level is "Task" unless otherwise stated.

The left-most state in each FSM is a dummy start state; the states that are double circles are states corresponding to the end of a two-turn sequence. By tracing from the start state to any one of the end states, it is possible to cover any of the selected two-turn pairs.

To keep the size of these FSMs at a minimum, we do not show turn sequences that occurred less frequently.

Statements

Assert The machine in figure 5 shows pairs of turns occurring more than 7 times, where at least one of the turns in each pair contains an utterance labeled with the *Assert* tag. We can see that a typical turn consists of a grounding utterance followed by one or more assertions, or of an assertion that is an answer, or of an action-directive that is an assertion. Looking at this machine, one might think that all turns contain only one or two utterances, but this is simply an artifact of our having to count only very frequently-occurring sequences of turns to reduce the size of the machine for presentation.

Reassert Turns containing reassertions frequently start with a grounding utterance (figure 6). It is interesting that *Holds* are so frequently followed by reassertions; one hopes that the form of the statement differs from the original assertion! It is also interesting that there are sequences of reassertions; both of these seem to run counter to Gricean maxims and to introduce redundancy into the dialog.

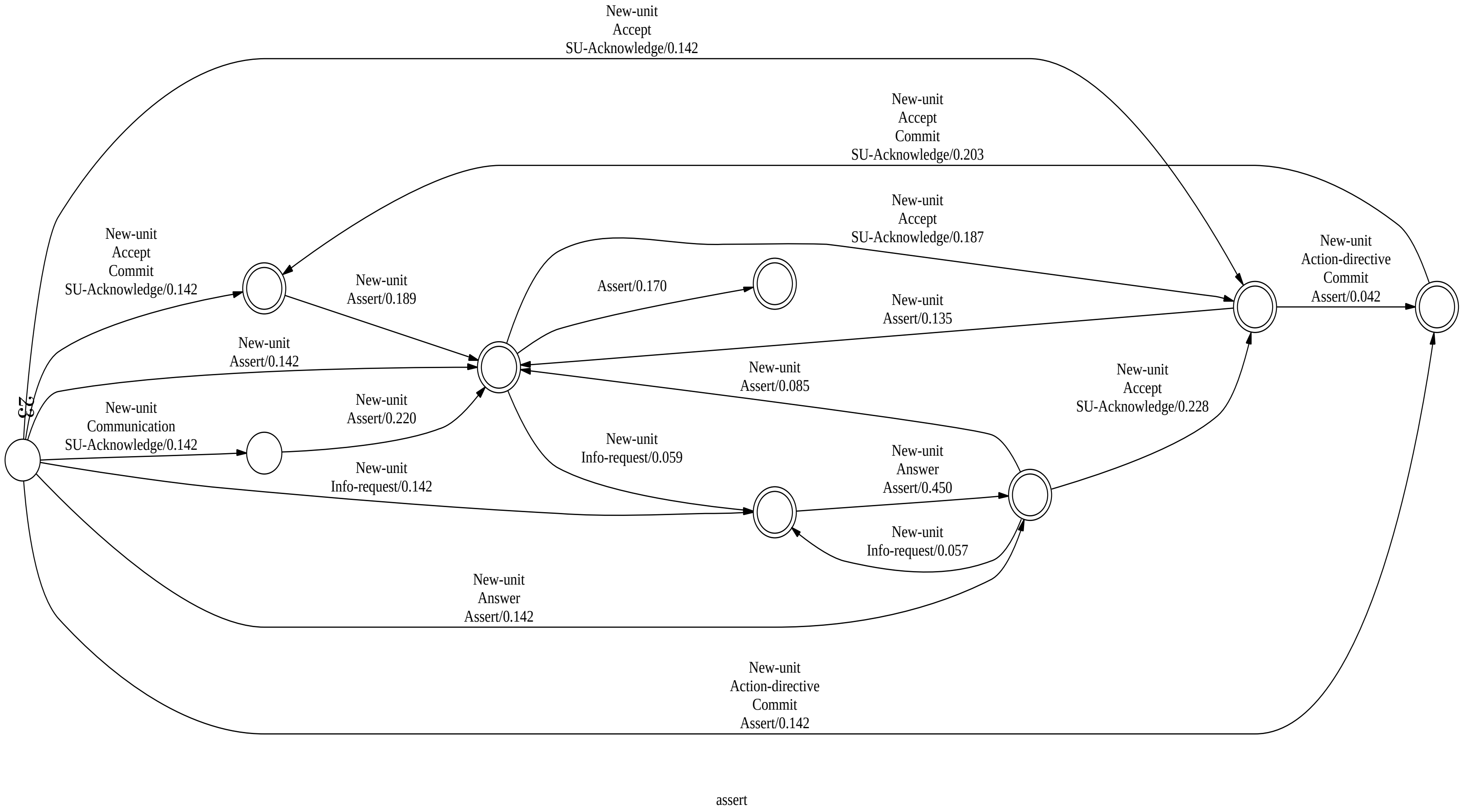

Figure 5: Example FSM: Context of *Assert*: turn pairs occurring 7 or more times

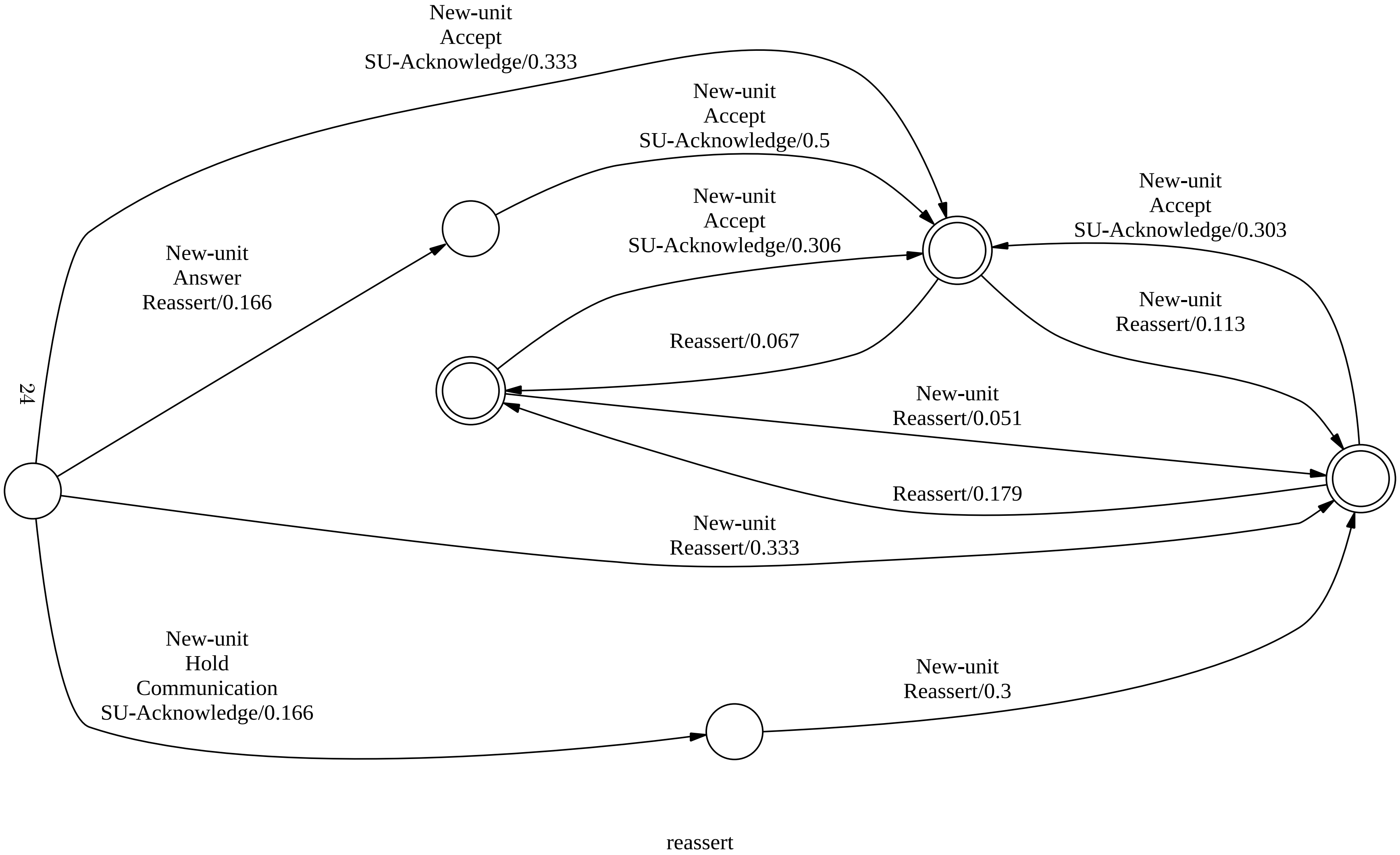

Figure 6: Example FSM: Context of *Reassert*: turn pairs occurring 3 or more times

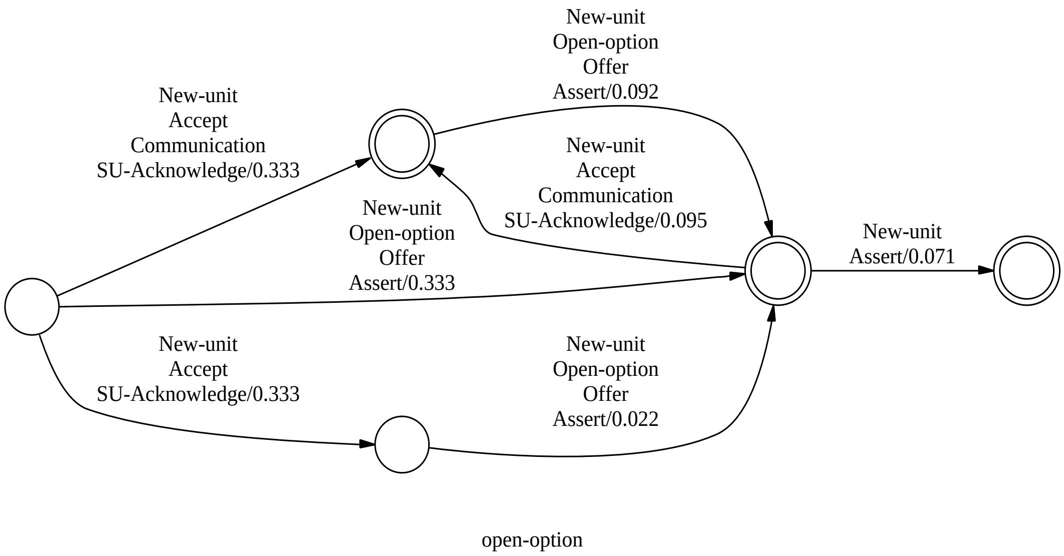

Figure 7: Example FSM: Context of *Open-option* and *Offer*: turn pairs occurring 3 or more times

Influence-on-listener and Influence-on-speaker

Action-directive Action-directives frequently occur in single-utterance turns. They are usually succeeded by an acceptance. The pattern then repeats.

From this, we can see that a conversational agent participating in planning dialogs may choose not to give a whole plan in one turn. Rather, it may try to divide up the plan description into several turns, each describing one action to be performed. This way, it is assured along the way that the other participant is understanding and agreeing with the proposed plan. By contrast, in text tutoring or explanation dialogs, the system can plan extended contributions. Part of the difference may be that in spoken dialog, there is no chance for the hearer to go back and re-hear, as a reader may read something several times. Another difference is that these dialogs are more mixed-initiative than in tutoring or explanation-generation systems, so more fluid interactions take place.

Commit The machine for the *Commit* tag is essentially the same as that for the *Action-directive* tag, because of the close coupling between action-directives and commits that we can see in figure 8. One would, perhaps, expect to see *Commits* following *Open-options*, but this does not seem to have occurred all that frequently. Typically, an *Open-option* is either challenged, or simply acknowledged. If it is just acknowledged, then at some later point it is repeated as an *Action-directive*.

Open-option and Offer The machines for these two tags are identical because, as we saw above, these two tags always co-occurred. We can see the same pattern as with the *Action-directive* tag: a proposal followed by an acknowledgment. The difference is that with an *Action-directive*, there is a much stronger obligation to express (dis-)agreement with the proposal; with *Open-option*, a simple acknowledgment is most-frequently seen.

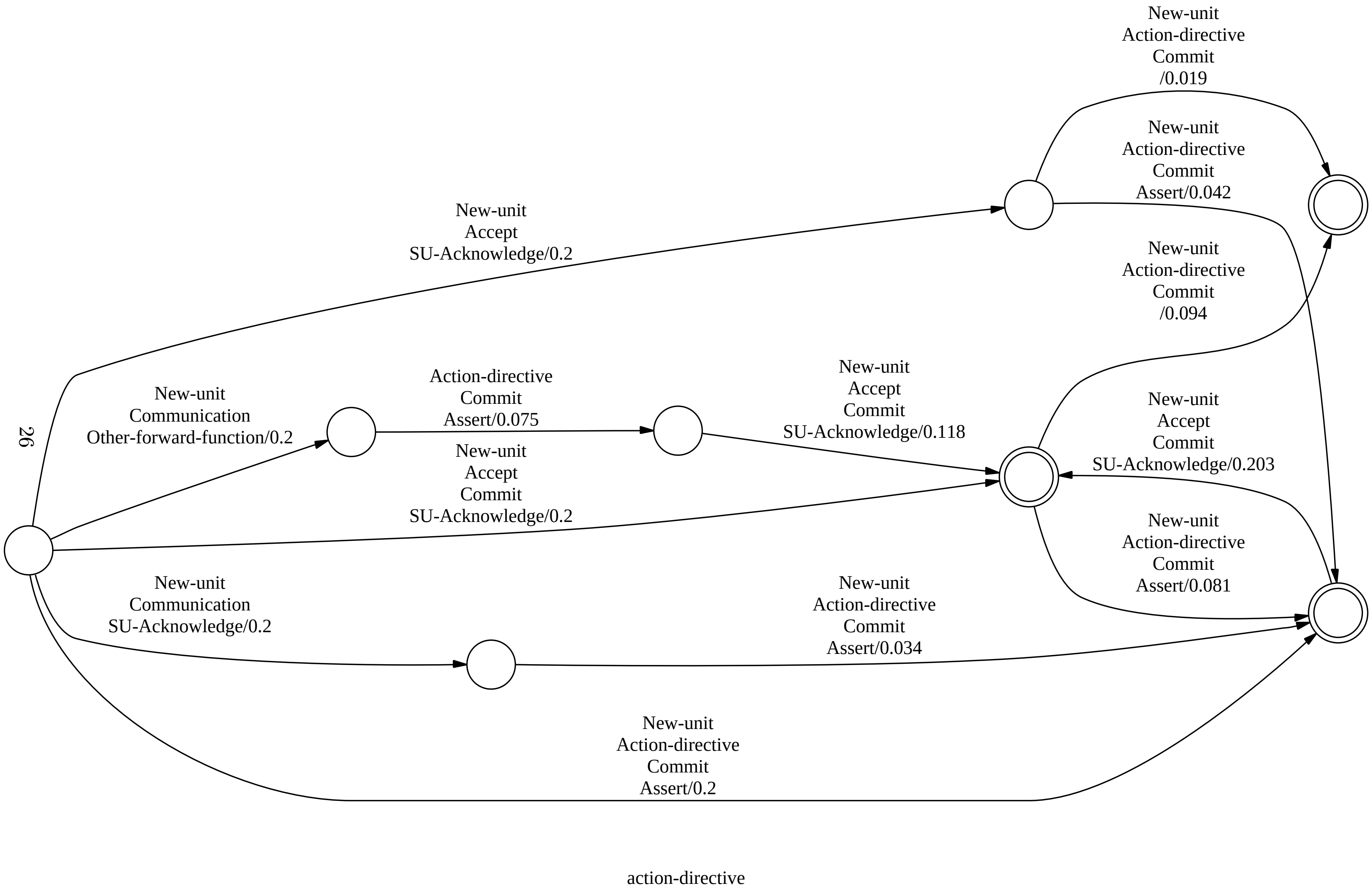

Figure 8: Example FSM: Context of *Action-directive*: turn pairs occurring 3 or more times

Info-request

Info-requests can follow statements, answers, or prompts. When they follow statements they can be requests for additional information or clarification questions. Following answers they tend to be clarification questions. When a speaker has to prompt the hearer to take the turn, the hearer probably needs to solve a problem or lacks understanding, and so we get questions there.

There can be sequences of questions; these occur more frequently than sequences of commands. This seems very counter-intuitive; however, often the second question is simply a restatement of the first or an elaboration on the first, e.g.:

Example 16

A	Where's Rochester General by the way
A	Is it [SIL] the [SIL] one right near the loop [SIL]

It might also be a prompt, as in:

Example 17

A	How close are Highland [SIL] and three eighty three two fifty two
A	Will you look at this

Not surprisingly, info-requests are typically followed by an answer from the other participant. In these dialogs, there were only a few times when the other participant could not make some kind of answer.

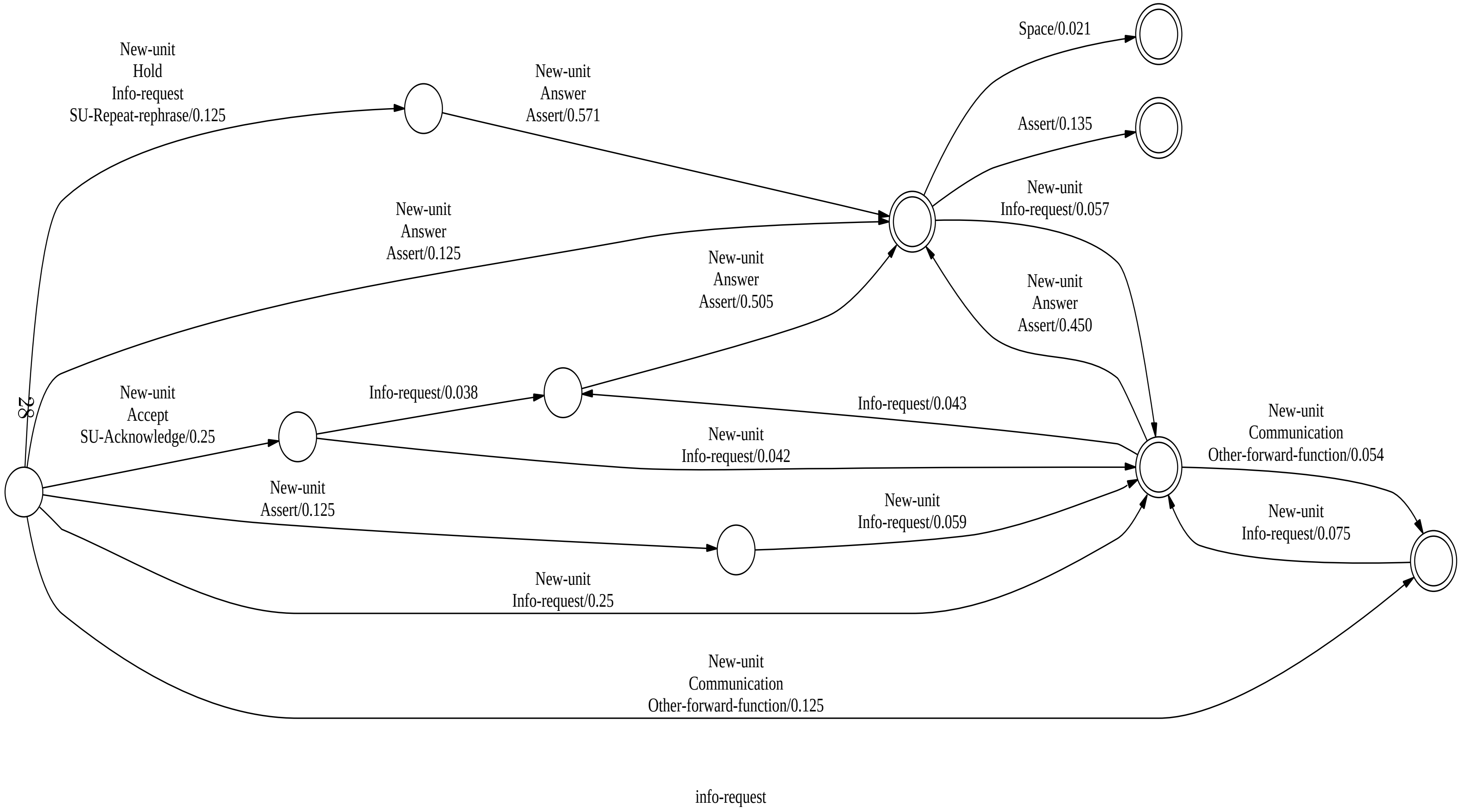

Figure 9: Example FSM: Context of *Info-request*: turn pairs occurring 3 or more times

Other-forward-function

Most of these utterances occurred at the start of the turn. They include turn-taking utterances such as “Well”, “Uh” and “Okay so”. Also included are turn-keeping and turn-passing utterances such as “Just a second”. Probably some utterances that are really prompts, e.g. “Right?”, were marked *Info-request* rather than *Other-forward-function*. The *Other-forward-function* tag was probably not marked enough, because often an “Okay” or “Alright” at the beginning of a turn was marked as part of the first utterance rather than as a separate utterance.

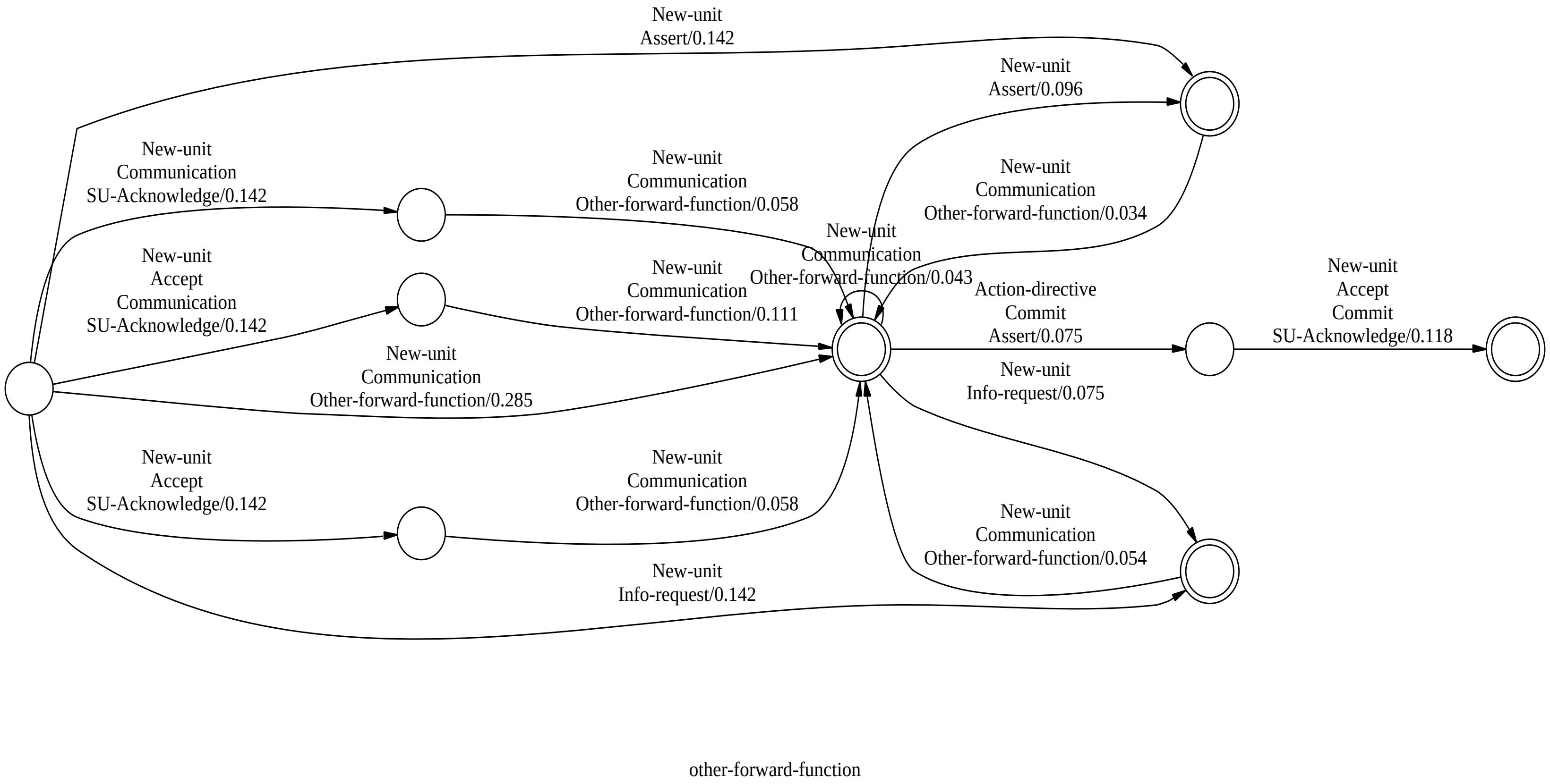

Figure 10: Example FSM: Context of *Other-forward-function*: turn pairs occurring 3 or more times

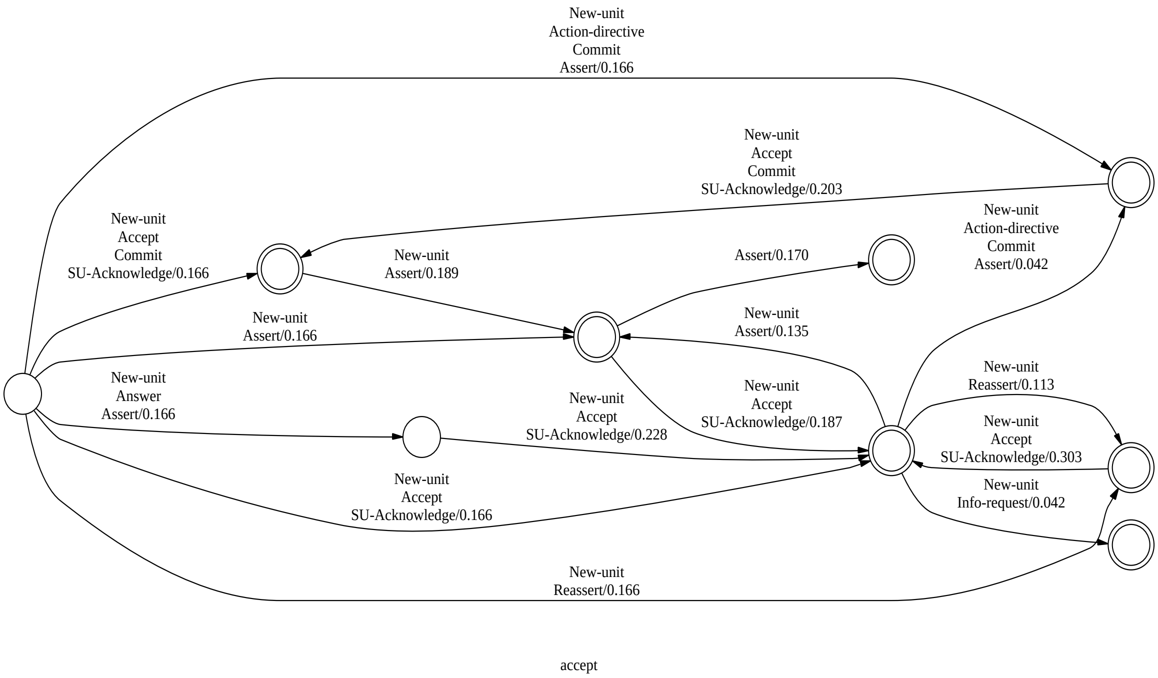

Figure 11: Example FSM: Context of *Accept*: turn pairs occurring 6 or more times

Agreement

Accept We can see from figure 11 that an *Accept* can follow many other types of act: answers, statements, or commands. *Accepts* most frequently occur in single-utterance turns, but may be followed by a statement. This tag does not frequently follow or come before an info-request.

The reason more *Accepts* do not occur in multiple-utterance turns is that, while these dialogs are mixed-initiative, the initiative is fairly stable. That is, the initiative does not generally change with each turn. The participant who does not have the initiative at the moment generally plays a fairly passive role, limited to asking questions, acknowledging and accepting the other participant's utterances, and waiting for a prompt to take the initiative. It may be true that a conversational agent would want to take the initiative at unusual times, such as when it gets a report that the human might be interested in; then, the system should pause between acknowledgment and taking of initiative so that the human does not think that the new topic is in some way related to the old.

Hold While this tag did not occur all that frequently, the part of the dialog model relating to it was so interesting that we include it anyway.

The *Hold* tag may be preceded by a wide variety of utterances, none of which occurred very frequently in our corpus. If it occurs with an *SU-Acknowledgment*, then it seems to indicate uncertainty on the part of the speaker. Then other participant repeats their assertion. At least in these dialogs, the other participant did not seem to make much of an attempt to justify their assertion; this is not what one would expect and should be further investigated.

If the *Hold* tag occurs with an info-request, it is a request for clarification, and the other participant answers. Here, the uncertainty may be expressed in two ways: by repeating (part of) the other participant's utterance with a questioning intonation, or by acknowledging it. An example of the first is the following:

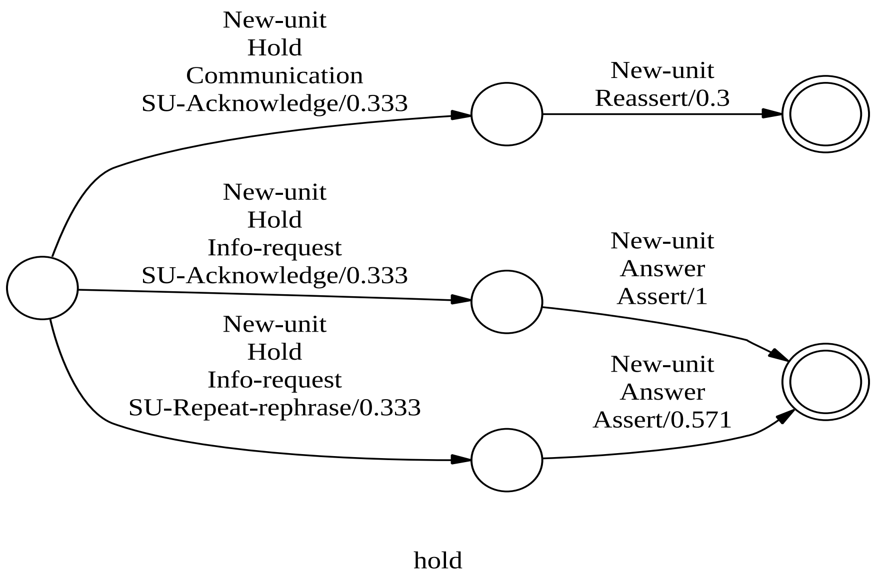

Figure 12: Example FSM: Context of *Hold*: turn pairs occurring 2 or more times

Example 18

A	Okay [SIL] so there are still thirty people [SIL] um [SIL] [click] [SIL] a [SIL] at [SIL] three ninety [SIL] that we need to [SIL] take care of [SIL]
B	Thirty people at three ninety?

An example of the second is:

Example 19

A	[SIL] Twenty-nine people
B	Um

Usually, a repetition is a strong indication of understanding; however, it can be used equally well to express lack of understanding.

Usually, in TRIPS the system tries to give as clear an indication of what type of misunderstanding is occurring as possible (reference failure, parse failure, failure to hear). In these dialogs, there is much more ambiguity, and the participants seem to handle it quite well. However, perhaps humans trust each other to be well-intentioned or intelligent more than they trust computer systems.

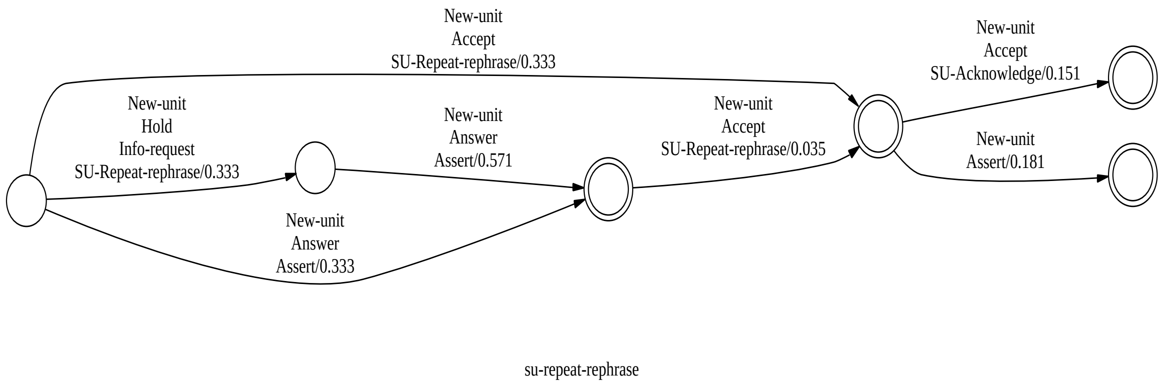

Figure 13: Example FSM: Context of *SU-Repeat-rephrase*: turn pairs occurring 3 or more times

Understanding

SU-Acknowledge This tag behaves a lot like the *Accept* tag: it follows all kinds of statements: assertions, action-directives and answers. It frequently occurs in single-utterance turns.

Because this tag occurs so frequently, in order to make the FSM readable we had to cut out a lot of sequences that actually occurred quite often; for instance, there are many two-utterance turns consisting of an acknowledgment followed by some sort of statement, but the combinations that go with the statement tag (influence-on-listener, influence-on-speaker, etc.) mean that no one type of statement occurred more than 7 times in this corpus. One conclusion we can draw, however, is that a conversational agent capable of participating in natural dialog will need to plan this type of act.

SU-Repeat-rephrase As we discussed briefly above, this tag is used for utterances that perform two very different functions. First, it is used when the speaker is unsure, to highlight the point where a lack of understanding occurred (although it remains unclear whether the speaker is checking for correct hearing or indicating a lack of comprehension). Second, it is used to emphasize a strong degree of understanding, for instance when answers to questions are repeated. Sometimes a repetition will itself be acknowledged. We can see that all of these happened in these dialogs (figure 13).

It is still unclear in what context a conversational agent might choose simply to acknowledge, and when it might choose instead to complete the previous speaker's thought or to repeat it. A system might take too long to produce words or might be too polite to use completions very often. However, repetitions might be effectively used to indicate the point of a reference failure.

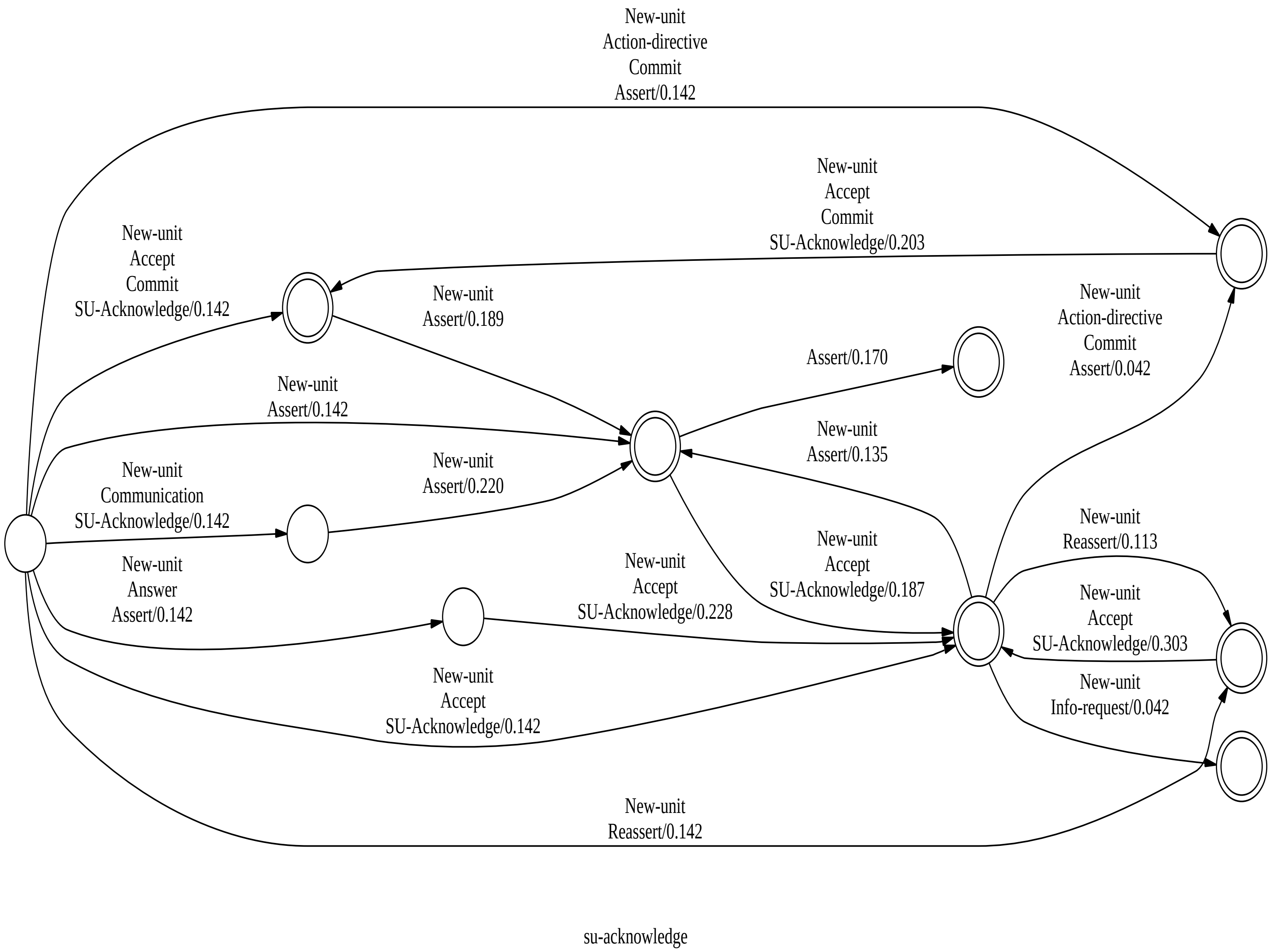

Figure 14: Example FSM: Context of *SU-Acknowledge*: turn pairs occurring 7 or more times

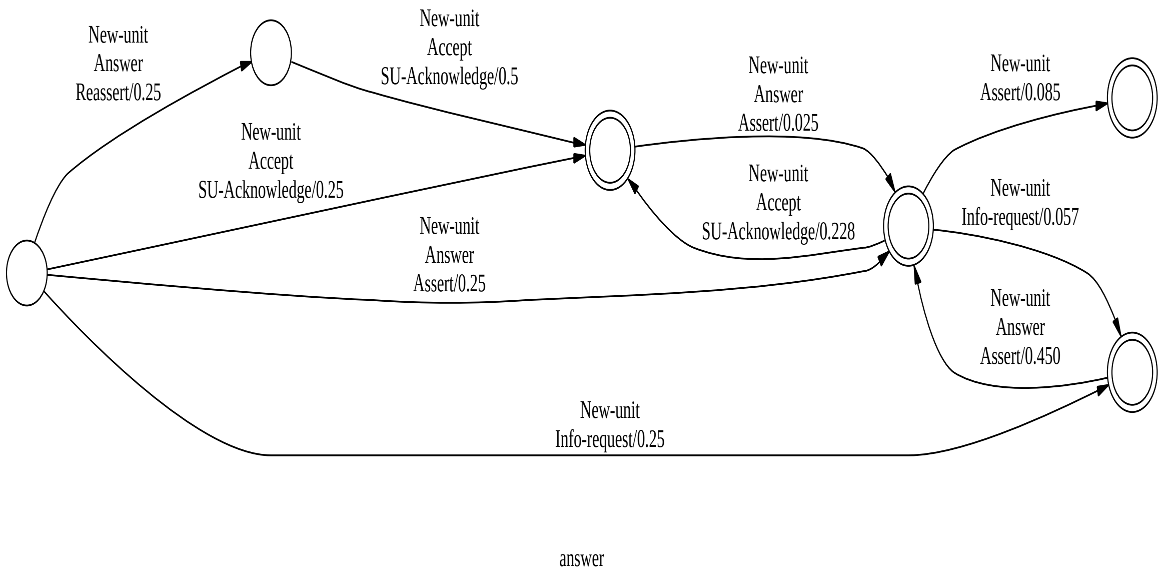

Figure 15: Example FSM: Context of *Answer*

Answer

The interesting thing about answers in this domain is that they are very often followed by an acknowledgment of the answer (figure 15):

Example 20

A	[SIL] [lipsmack] [SIL] So bus B is going to have to make a second trip
B	Yeah
A	Okay

This may be because answers to questions are important; lack of information can hold up planning, and conflicting information can both hold up planning and confuse dialog participants. The acknowledgment also serves the purpose of signaling a pop out of the question sub-dialog. A conversational agent should therefore not plan to combine an answer with some other utterance, but should wait for the acknowledgment.

There were, however, times when an answer was followed by another utterance by the same speaker. This was usually an elaboration of or motivation of the answer, e.g.:

Example 21

A	That's all bus [SIL] B s has to do?
B	Yeah
B	Because it has to travel a pretty far [SIL] distance [SIL] + to [SIL] two sixty + one so

Monroe	Switchboard-DAMSL
Statement (50.9%)	Statement-non-opinion (36%), Statement-opinion (13%)
Action-directive (7.9%)	Action-directive (.4%)
Open-option, Commit, Offer (21.5% total)	Open-option, Commit, Offer (.1%)
Info-request (9.8%)	(Declarative) Yes-No-Question (3%), (Declarative) Wh-Question (1.1%), Open-Question (.3%), Rhetorical-Questions (.2%)
Other-forward-function (5.7%)	Tag-Question (<.1%), Apology (<.1%), Thanking (<.1%), Hedge (1%), Other (1%)
Explicit-performative (0.3%)	none
Exclamation (1.2%)	none
Opening (0.3%)	Conventional-opening (.1%)
Closing (0.3%)	Conventional-closing (1%)
Accept (22.8%)	Agree/Accept (5%)
Accept-part (0.2%)	none
Reject (0.5%)	Reject (.2%)
Reject-part (0.2%)	none
Maybe (0.2%)	Maybe/Accept-part (<.1%)
Hold (1.2%)	Hold (.3%)
Signal-non-understanding (0.5%)	Signal-non-understanding (.1%)
Correct-misspeaking (0.3%)	none
SU-Acknowledge (24.2%)	Acknowledge (19%), Response-acknowledgement (1%)
SU-Repeat-rephrase (3.3%)	Summarize/reformulate (.5%), Repeat-phrase (.3%)
SU-Completion (0.7%)	Collaborative completion (.4%)
Answer (8.3%)	Various types of answer totaling 2.7%

Table 19: Comparison of Monroe and Switchboard-DAMSL tag frequencies

7 Comparison with other results

For most of the other dialog annotation projects, there are not publicly available results of the kind we report here. However, the Switchboard-DAMSL project ([14]) gives the counts of tags for the slightly different dialog acts set they used.

In the Switchboard-DAMSL project, 205,000 utterances from 1155 non-task-oriented dialogs were labeled. 197,489 of these were used as training-set utterances; these are the utterances for which statistics are given. Because the corpus was not task-oriented, there are some significant differences in the frequencies of various tags. For instance, there are many more action-directives in the Monroe corpus; there are also more questions. However, both these sets of results illustrate the frequency of grounding acts in spoken dialog.

8 Future work

8.1 More detailed annotation

We have, at this point, data annotated for illocutionary acts and grounding acts. In the near future, we plan to annotate for turn-taking acts. We are currently exploring ways to code dialog structure; previous work has shown that taking this into account can improve the accuracy of probabilistic dialog models [19]. We are also writing an annotation manual for argumentation acts.

8.2 Bootstrapping for other annotations

There are some annotation schemes that require (or can be made to require) the use of exactly the kinds of information a DAMSL annotation provides, specifically, information about illocutionary acts and relationships between utterances. For instance, Chu-Carroll and Brown's scheme for marking or predicting initiative shifts in dialog [5] depends on the annotator being able to tell the type of act and dialog level of an utterance.

We can have the computer automatically produce an initial annotation for initiative shifts, given a DAMSL annotation of a dialog. While not complete, this initial annotation provides quite a bit of assistance for the human annotator.

8.3 Building dialog models

In the field of dialog interpretation, there have been several approaches using speech-act annotated data to build dialog models for speech-act prediction and language modeling (e.g. [20], [15], [19] [14]). None of these researchers were concerned with using speech-act annotated data to aid content planning or generation. Furthermore, all except [14] used single-layer annotation schemes. We plan to experiment with building statistical dialog models of this nature. At the least, they furnish probabilities to assign to rules in a generation component of a dialog system. They may even be useful as part of the generation component themselves.

8.4 Theory development and implementation

We are using our annotated data as we plan a system for generation in the context of free-flowing dialog([21]). This system is in the early stages of implementation; so far, the data has been used primarily to help us think about the architecture and construct simple human-computer dialogs. However, we plan to make more use of our data as implementation progresses.

9 Conclusions

The Monroe corpus is a mixed-initiative, task-oriented corpus in a complex domain that provides a good resource for examining different dialog behaviors. DAMSL provides a clear annotation scheme that can be used by relatively inexperienced annotators. The resulting annotated data is useful for describing dialog behaviors, for bootstrapping other annotations, and for designing and building a generation component for free-flowing spoken dialog.

References

[1] J. Allen and M. Core. Draft of DAMSL: Dialog Act Markup in Several Layers, 1997. available at: http://www.cs.rochester.edu/research/trains/annotation/RevisedManual/RevisedManual.html.

[2] Anne H. Anderson, Miles Bader, Ellen Curman Bard, Elizabeth Boyle, Gwyneth Doherty, Simon Garrod, Stephen Isard, Jacqueline Kowtko, Jan McAllister, Jim Miller, Catherine Sotillo, Henry Thompson, and Regina Weinert. The HCRC Maptask corpus. *Language and Speech*, 34(4):351–366, 1991.

[3] J. Carletta. Assessing agreement on classification tasks: the kappa statistic. *Computational Linguistics*, 22(2):249–254, 1996.

[4] J. C. Carletta, A. Isard, S. Isard, J. Kowtko, G. Doherty-Sneddon, and A. Anderson. The reliability of a dialogue structure coding scheme. *Computational Linguistics*, 23:13–31, 1997.

[5] J. Chu-Carroll and M. Brown. Initiative in collaborative interactions – its cues and effects. In *Proceedings of the AAAI Spring Symposium on Computational Models for Mixed Initiative Interaction*, March 1997.

[6] H. Clark and E. Schaeffer. Contributing to discourse. *Cognitive Science*, pages 259–294, 1989.

[7] M. Core. Addendum to coding dialogs with the DAMSL annotation scheme. ftp://ftp.cs.rochester.edu/pub/papers/ai/97.Core-Allen.AAAI2.ps.gz, 1997.

[8] M. Core. DAMSL Annotation Tool, version 1.10, 1997. available at: http://www.cs.rochester.edu/research/trains/annotation.

[9] M. Core and J. Allen. Coding dialogs with the DAMSL annotation scheme. In *AAAI Fall Symposium on Communicative Action in Humans and Machines*, pages 28–35, November 1997.

[10] B. Di Eugenio, P. Jordan, J. Moore, and R. Thomason. An empirical investigation of proposals in collaborative dialogues. In *Proceedings of the 17th International Conference on Computational Linguistics and the 36th Annual Meeting of the Association for Computational Linguistics (COLING-ACL'98)*, 1998.

[11] B. Di Eugenio, P. Jordan, and L. Pylkkänen. The COCONUT project: dialogue annotation manual. Technical report, University of Pittsburgh, 1998.

[12] P. Heeman and J. Allen. The TRAINS 93 dialogs. Technical Report Trains TN 94-2, Computer Science Dept., U. Rochester, March 1995.

[13] Peter A. Heeman and James F. Allen. Dialogue transcription tools. Technical report, Computer Science Dept., U. Rochester, August 1994; revised March 1995.

[14] Daniel Jurafsky, Rebecca Bates, Noah Coccaro, Rachel Martin, Marie Meteer, Klaus Ries, Elizabeth Shriberg, Andreas Stolcke, Paul Taylor, and Carol Van Ess-Dykema. Switchboard discourse language modeling project report. Technical Report Research Note 30, Center for Speech and Language Processing, Johns Hopkins University, Baltimore, MD, 1998.

[15] K. Kita, Y. Fukui, M. Nagata, and T. Morimoto. Automatic acquisition of probabilistic dialogue models. In *Proceedings ICSLP 96*, volume 1, pages 196–199, 1996.

[16] K. McKeown. *Text Generation: Using Discourse Strategies and Focus Constraints to Generate Natural Language Text.* Cambridge University Press, Cambridge, 1985.

[17] J. D. Moore and C. L. Paris. Planning text for advisory dialogues: Capturing intentional and rhetorical information. *Computational Linguistics*, 19(4):651–695, 1993.

[18] L. Pineda. Personal correspondence.

[19] M. Poesio and A. Mikheev. The predictive power of game structure in dialogue act recognition: Experimental results using maximum entropy estimation. In *Proceedings of ICSLP-98*, 1998.

[20] N. Reithinger. Some experiments in speech act prediction. In *Empirical Methods in Discourse Interpretation and Generation. Papers from the 1995 AAAI Symposium*, number TR SS-95-06, pages 126–131, Menlo Park, CA, 1995. AAAI Press.

[21] A. Stent. Content planning and generation in continuous-speech spoken dialog systems. In *to appear in the Proceedings of the KI'99 workshop, "May I Speak Freely?"*, 1999.

[22] D. Traum and E. Hinkelman. Conversation acts in task-oriented spoken dialogue. *Computational Intelligence*, 8(3):575–599, 1992.

A Notes on Dialogs

Dialog	Subjects	Task	Length (sec.)	Completed initial task?	Extra task assigned?	Completed extra task?
1	S1R, S1B	1	459	yes	yes	yes
2	S1B, S1R	2	706	yes	yes	yes
3	S2B, S2R	3	560	yes	yes	no
4	S2R, S2B	4	1072	yes	yes	yes
5	S3B, S3R	5	2342	yes	no	
6	S3R, S3B	6	1623	yes	no	
7	S4R, S4B	8	1711	yes	no	
8	S4B, S4R	7	1435	yes	no	
9	S5B, S5R	5	2433	yes	no	
10	S5R, S5B	1	584	yes	no	
11	S6R, S6B	6	974	yes	no	
12	S6B, S6R	2	418	yes	no	
13	S7B, S7R	7	2144	yes	no	
14	S7R, S7B	3	771	yes	yes	yes
15	S8R, S8B	8	592	yes	yes	yes
16	S8B, S8R	4	1017	yes	yes	yes
17	S9R, S9B	8	891	yes	yes	no
18	S9B, S9R	1	1230	yes	yes	yes
19	S10B, S10R	4	1603	yes	yes	yes
20	S10R, S10B	5	1372	yes	no	

Dialog 1: S1R had trouble hearing S1B, who spoke with the microphone a little far away from his face. Tape starts a little into the dialog.
Dialog 7: S4B says "three" like "tree". S4R invented vehicles.
Dialog 8: S4B says "three" like "tree".
Dialog 13: Both speakers have colds.
Dialog 14: Phone rings midway through dialog. Both speakers have colds.
Dialog 15: S8R says "ambulance" in an unusual way.
Dialog 16: S8R says "ambulance" in an unusual way.
Dialog 17: S9R puts pen in mouth about 3/4 way through dialog, then removes it after a few utterances.
Dialog 18: S9R either chews his fingers or puts things in his mouth.
Dialog 19: S10B has a cold.
Dialog 20: S10B has a cold.

B Tasks

The eight tasks we used are given on the following pages.

1. There are three school buses, two at the Pittsford fire station and

one at the Irondequoit police station. There is also one helicopter at Strong Memorial hospital. There are two medivac crews, one at the Chili fire station and one at the fire station in Penfield. Neither has transport of their own.

Because of an approaching storm, houses along the lake are being evacuated. There are 70 people at the end of route 261, 35 at the end of highway 390N and 30 at the end of highway 590N. One person at the end of route 261 is ill (needs immediate transportation to a hospital) and one at the end of highway 590N is handicapped and so cannot ride the bus.

Your task is to get all the people to safety as fast as possible. Flooding is expected and will reach all lakeshore areas in 3 hours. There is a shelter at the Irondequoit police station to which people can be moved. The ill person must be transported to a hospital.

Transport-type	Number	Location
bus	2	Pittsford fire station
bus	1	Irondequoit police station
helicopter	1	Strong Memorial hosp.

Transports

Number	Status	Location
69	healthy	Rte. 261
1	ill	Rte. 261
35	healthy	Hwy. 390N
29	healthy	Hwy. 590N
1	handicapped	Hwy. 590N

Locations of people

2. There has been a bomb attack at the airport, and as a result there are 6 injured people there, 3 of whom are on stretchers. In addition, there is an individual with a heart attack at Marketplace Mall, and one with a broken leg in Highland park. Your task is to transport all patients to hospitals as soon as possible, considering their stability and feasible transports.

There are two ambulances and one helicopter at Strong Memorial hospital. Emergency crews for all transports are also at Strong.

The helicopter can land only at the airport and Strong. Patients can be transported to St. Mary's or to Strong, but St. Mary's doesn't have a cardiac care center.

Number	Status	Location
3	walking wounded	Roch. Int. airport
3	stretcher-bound	Roch. Int. airport
1	stretcher-bound	Marketplace Mall
1	walking wounded	Highland park

Locations of people

Transport	Number	Location
ambulance	2	Strong Memorial hosp.
helicopter	1	Strong Memorial hosp.

Transports

3. There are two road crews, one each at the airport and Gates Center. There are two electrical crews at the main RGE plant on route 33 at route 33A. All crews have trucks and equipment.

As a result of an ice storm, roads are out at the highway 490 bridge over the river, Beahan Rd. (near the airport), and route 383 at route 252A. Major electric lines are down at route 383 at route 252A, Highland Ave, and Beahan Rd. At route 383 the line is lying across the road.

Your task is to make a plan to fix all roads and major electric lines as soon as possible.

Roads must be fixed before electric lines can be fixed, but if a line is lying across the road the electric power must be shut off by an electric crew at the two nearest intersections before the road is fixed (this takes 20 minutes). It takes 2 hours to fix an electric line and 4 to fix a road.

Location	Road out	Electric lines down	Line across road
490 bridge	yes	no	no
Beahan Rd.	yes	yes	no
Rte. 383 and Rte. 252A	yes	yes	yes
Highland Ave.	no	yes	no

Locations of ice storm problems

Crew	Number	Repair time	Location
road	1	4 hours	Roch. Int. airport
road	1	4 hours	Gates Center
electric	2	2 hours	Rte. 33 and Rte. 33A

Crews available

4. As a result of an ice storm, roads are out at the Elmwood bridge, Thurston Rd, route 383 at Brooks Ave., and highway 490 at the inner loop (at Thurston Rd. and the Elmwood bridge digging will be required). Major electric lines are down at route 383 at Brooks, and route 31 at highway 390. At route 31 the line is lying across the road.

There is an individual on oxygen at route 383 just past route 252A; a generator must reach this person before their emergency battery expires in 3 hours. There is an individual with pneumonia at Main St. and East Ave. (route 33A and route 96) who needs to be taken to hospital as soon as possible.

Your task is to make a plan to take care of the individuals in need and get the roads fixed and the major electrical lines repaired as soon as possible.

There are two road crews, one at the municipal building on Main St. and East Ave., and one at the airport. There are three electrical crews, all at the main RGE plant on route 33 at route 33A. There is one mobile emergency generator, currently at Strong Memorial hospital, and one truck with equipment for digging, currently at the main RGE plant. The crews have all other necessary equipment. There is an ambulance (with crew) at Strong.

Roads must be fixed before electric lines can be fixed, but if a line is lying across the road the electric power must be shut off by an electric crew at the two nearest intersections before the road is fixed (this takes 20 minutes). It takes 2 hours to fix a section of road and 1 1/2 hours to fix a power line. To dig takes a maximum of 1 hour.

Location	Road out	Electric line down	Other
Elmwood bridge	yes	no	digging required
Thurston Rd.	yes	no	digging required
Rte 383 and Brooks Ave.	yes	yes	person near location
Hwy 490 and inner loop	yes	no	
Rte. 31 and Hwy 390	no	yes	line across road
Main St. and East Ave.	no	no	person at location

Locations of ice storm problems

Type	Number	Repair time	Location
Road crew	1	2 hrs.	Main St. and East Ave.
Road crew	1	2 hrs.	Roch. Int. airport
Electric crew	3	1 1/2 hrs.	Rte. 33 and Rte. 33A
Generator	1	-	Strong Memorial hosp.
Digging truck	1	1 hr.	Rte. 33 and Rte. 33A
Ambulance	1	-	Strong Memorial hosp.

Crews and equipment

5. There has been a major snow storm, resulting in burst water mains and downed power lines. The city has a routine schedule for plowing the roads; you have three snow plows for emergency use to clear roads for repairs. You also have two road crews to fix water mains and three

electric crews to fix downed power lines. The snow plows and road crews are at the main municipal garage on Main St. and East Ave. (the intersection of route 33A and route 96); the electric crews are at RGE on route 33 at route 33A. There is one machine for digging underground, also at RGE.

The locations of major downed power lines are: route 252 at the river, East Ave. at highway 590, route 286 at highway 590, Monroe Ave. (route 31) at Highland Ave., and route 65 at route 253. The locations of broken water mains are: highway 490 at the inner loop, route 253 at route 65, highway 386 at route 104, East Ave. at Culver Rd., and the airport (which has priority). At route 252 and Highland Ave. the lines are on the road.

There are three people in wheelchairs dependent on electricity, two at East Ave. between highway 590 and N. Goodman, and one at route 383 next to route 252 (route 252 electric line); if electricity can not be restored to them in 3 hours an ambulance must be sent to fetch them. There is an ambulance (with crew) at Strong.

Roads must be fixed before electric lines can be fixed, but if a line is lying across the road the electric power must be shut off by an electric crew at the two nearest intersections before the road is fixed (this takes 20 minutes). A road must be plowed before any service vehicles can proceed along it. You can assume all major roads (roads with numbers) are already plowed. It takes 3 hours to fix a water pipe (including digging), 1 hour to fix a power line, and 10 minutes to plow 1 mile of road.

Your task is to get the power lines and water mains fixed, taking into account local priorities.

Location	Water main	Power line	Other
Rte. 386 and Rte. 104	yes	no	
East Ave. and Culver Rd.	yes	no	
Rte. 252 at river	no	yes	person at location, line across road
East Ave. and Hwy. 590	no	yes	2 people at location
Rte. 286 and Hwy. 590	no	yes	
Monroe Ave. and Highland Ave.	no	yes	line across road
Rte. 65 and Rte. 253	yes	yes	
Hwy. 490 at inner loop	yes	no	
Roch. Int. airport	yes	no	

Locations of disturbances

Equipment	Number	Repair time	Location
Snow plow	3	10min/mile	West Main St. and East Ave.
Road crew	2	3 hrs.	West Main St. and East Ave.
Electric crew	3	1 hr.	Rte. 33 and Rte. 33A
Digging machine	1	-	Rte. 33 and Rte. 33A
Ambulance	1	-	Strong Memorial hosp.

Crews and equipment

6. There has been an ice storm which has left power lines down and roads blocked by trees. Your job is to get the roads cleared and the power lines fixed as fast as possible, bearing in mind the importance of roads and locations of hospitals.

There are power lines down on Elmwood Ave. at S. Clinton, highway 590 at East Ave. (route 96), highway 390 at Mt. Hope Ave., and Genesee St. There are trees down at Genesee St., Mt. Hope Ave. at Highland Park, route 252 at Marketplace Mall, route 383 at route 252A, and highway 590 at East Ave.

Downed trees must be removed before power lines can be fixed. It takes about 45 minutes to fix a power line.

If a road is next to a park, the tree can be pulled into the park (1/2 hour); if it is not, the tree must be cut up (2 hours).

There are two crews available to fix downed power lines or remove trees from the road. They have at their disposal one truck that can be used to fix lines or pull a tree, one truck that can carry large chunks of cut-up tree only, and one truck for the use of fixing power lines only. Each crew has four people; all crews and equipment are at RGE (at the intersection of routes 33 and 33A). Another crew will be

arriving at RGE in an hour from Irondequoit, with another truck that can be used to fix lines.

Location	Power line down	Tree down
Elmwood Ave. and Clinton Ave.	yes	no
Hwy. 590 and East Ave.	yes	yes
Hwy. 390 and Mt. Hope Ave.	yes	no
Genesee St.	yes	yes
Mt. Hope Ave. at Highland Park	no	yes
Rte. 252 at Marketplace Mall	no	yes
Rte. 383 and Rte. 252A	no	yes

Ice storm disturbances

Type of repair	Repair time
Fix line	45 min
Tow tree	30 min
Chop and carry pieces	2 hrs

Repair times

7. A tornado has left Rochester in a panic. There are power lines down on highway 390 at route 104, route 31 at Mt. Read Blvd., Elmwood Ave. at Strong Memorial hospital, route 33A at Howard Rd., and route 15 at route 252. 100 residents of a mobile home park at route 261 and route 31 need to be moved to temporary shelter at Gates Center; 2 more are injured (one is on a stretcher) and need to be transported to hospital. There are trees down at route 31 at Mount Read. Blvd. and route 15 at route 252. Finally, there is a Brinks' armoured car at highway 490 at Mt. Read Blvd. that needs towing, having been flipped in the tornado.

You have two buses at your disposal, both currently at the municipal building on East Ave. and Main St. (route 96 and route 33A). Strong Memorial hospital has a helicopter and 2 ambulances, all with crews.

There are two crews available to fix downed power lines or remove trees from the road. They have at their disposal one truck that can be used to fix lines or tow heavy items (cars or trees), one truck that can carry large chunks of cut-up tree, and one truck for the use of fixing power lines only. Each crew has four people; all crews and equipment are at RGE (at the intersection of routes 33 and 33A). Another crew will be arriving at RGE in an hour from Irondequoit.

It takes 20 minutes to tow a tree off the road and 1 1/2 hours to chop a tree up. A power line can be fixed in 1 hour.

Your task is to arrange to fix these difficulties as fast as possible.

Location	Disturbance
Hwy. 390 and Rte. 104	power line down
Rte. 31 and Mt. Read Blvd.	power line down, tree down
Elmwood Ave. at Strong Mem. hosp.	power line down
Rte. 33A and Howard Rd.	power line down
Rte. 15 and Rte. 252	power line down tree down
Rte. 31 and Rte. 261	100 people
Rte. 31 and Rte. 261	2 injured people
Hwy. 490 and Mt. Read Blvd.	armored car

Disturbances

Equipment	Number	Location	Use
Bus	2	Rte. 96 and Rte. 33A	-
Helicopter	1	Strong Mem. hosp.	-
Ambulance	2	Strong Mem. hosp.	-
Utility crews	2/3	Rte. 33 and Rte. 33A	-
Truck	1	Rte. 33 and Rte. 33A	fix lines/tow
Truck	1	Rte. 33 and Rte. 33A	carry pieces
Truck	1	Rte. 33 and Rte. 33A	fix lines

Equipment and crews

Type of repair	Repair time
Fix line	1 hr
Tow tree	20 min
Chop and carry pieces	1.5 hrs

Repair times

8. Rochester is undergoing political turmoil. There is rioting in the downtown area on Lake Ave., South Clinton, and Rte. 31 on both sides of the inner loop; and in the east on E. Main St. (route 33A inside the inner loop), Culver Rd. and East Ave (route 96). Reports are that there are three injured on State St. at the inner loop; there are five on N. Goodman at East Main St. (2 on stretchers); and there are three on East Ave. at Culver (1 on a stretcher). A fire hydrant has been broken and is spraying water into the air on route 31 at the south side of the inner loop, and there is a fire on Lake Ave. at the inner loop.

Extra police units must reach the area before any emergency crews can. There are four units at the police station in Pittsford; a fire truck is waiting at the Pittsford fire station; the police station on Main St. and East Ave. (route 33A and route 96) has three police units, and there is 1 ambulance at St. Mary's hospital and 2 at Rochester General. A water crew (with equipment) is waiting at the municipal building at West Main St. and East Ave.

Make a plan to return order as fast as possible.

Location	rioting	injuries (walking/stretcher)	Other
Lake Ave.	yes	3/0	fire
South Clinton Ave.	yes		
Rte. 31 off 33A	yes		
Rte. 31 off Goodman St.	yes		broken hydrant
E. Main St.	yes	3/2	
Culver Rd.	yes	3/1	
East Ave.	yes		

Locations of difficulties

Force	Number	Location(s)
police unit	4	Pittsford P.S.
police unit	3	W. Main St.
fire truck	1	Pittsford F.S.
ambulance	1	St. Mary's hosp.
ambulance	2	Roch. Gen. hosp.
water crew	1	W. Main St. and East Ave.

Locations of emergency units

C Other information

The following information was given to all subjects.

Helpful information

Important locations	
Name	Access
St. Mary's hosp.	Genesee St. and W. Main St.
Rochester General hosp.	Portland Ave., Goodman Ave., Rte. 104
Monroe Community hosp.	Rte. 15A
Strong Memorial hosp.	Elmwood Ave.
Marketplace Mall	Rte. 252 and Rte. 15
Roch. Int. airport	Brooks Ave.
Highland park	Highland Ave. and Rte. 15

Speeds and capacities		
Vehicle	Speed (road/highway)	Capacity
Ambulance	45mph/70mph	3 seated or 2 stretchers
Police car	45mph/70mph	-
Fire truck	45mph/70mph	-
Utility truck	30mph/55mph	-
Helicopter	1/2 hr. in Monroe Co.	2 seated or 1 stretcher
Bus	30mph/55mph	30 seated
Snow plow	30mph/50mph	-

D Instructions

The instructions below were read aloud to all subjects at the appropriate times. The experimentor also told the subjects that she would only answer questions about the locations of items.

There are three parts to this study, and I am going to give you instructions for each part as we come to it.

D.1

Each of you has a paper with a description of a task on it, a map, and some other information on the table in front of you. I want you to pretend that you work at an emergency control center, and this is an emergency you have just been told about. You should make the best plan you can to solve the task.

You may write on the paper or the map. I will not answer questions, except about the locations of roads, so if there is some information you think is missing make a reasonable guess or estimate and write it down. I will give you ten minutes to work on the task. We do not expect you to complete your plan in that time, just try your best.

D.2

Now I want you to pretend that control of this emergency has been passed to another center, and you need to discuss your plan with the person at the other center who will be handling it.

I want you to pretend that you are the person at the other center. Your and your coworker should collaborate to finalize the plan. Feel free to ask questions and suggest corrections or alterations to improve the plan. When the discussion is over, you should summarize the plan for your coworker.

D.3

One of these extra problems was read orally to the subjects at the end of each dialog if there was time.

Now there is a new aspect to the task. I want you and your coworker to collaborate to modify your plan on the basis of this new information.

Task 1: 2 people at end of 590 become ill also. None of the ill people are on stretchers. OR One schoolbus has broken down, so instead two vans are being sent; each holds ten people.

Task 2: There are two more injured people at the airport, one on a stretcher and one not. OR There is a report of an injury on Jefferson at the river.

Task 3: There is a road out at Jefferson at the river. OR There is a road out on 15 between Elmwood and 15A.

Task 4: There is a person on Brooks Ave. with pneumonia who needs transport to the hospital as soon as possible. OR There is a road out on 15 between Elmwood and 15A.

Task 5: One of the snow plows broke down. OR There is a person on Browncroft Blvd who needs to be transported to the hospital.

Task 6: New trees down reported at East River Rd. before 252 and on 33A in West Chili. OR New tree down at Elmwood Ave. at South Clinton.

Task 7: 1 more injured at the mobile home park, on a stretcher. OR New tree down at Rte. 33A at Howard Rd.

Task 8: 1 more injured at Culver, on a stretcher. OR New fire on S. Clinton.

E Consent form

Consent Form
Data Collection and Analysis
Principal Investigator: James Allen
University of Rochester

We are investigating how people work together using language to solve problem-solving tasks. In order to identify problems for future study, we would like you, with another participant, to discuss solutions to some problems we will provide.

If you agree to participate, you will be given one or more problems to solve and asked to interact with another person using spoken English to share or complete your solution(s). You will be given a map and some tables to assist you. You will wear headphones and speak into a microphone. Your speech will be recorded, and your use of the map and tables will also be recorded using a video camera. This information will be saved for future study.

There is no risk to you in this study. The purpose is to collect data, not to test your problem-solving skills.

By participating in this study, you agree to allow us to analyze and distribute the recorded interactions, but you will never be personally identified as the speaker. Your participation is voluntary, and at any time you may withdraw with no change in circumstance. For your participation in this study, you will be paid $7.50.

____

Your name (printed)

________________________________ ______________

Signature Date

________________________________ ______________

Experimenter Date