USING THE PROCESS OF DISTILLING DIALOGUES TO UNDERSTAND DIALOGUE SYSTEMS

Staffan Larsson[1] *Lena Santamarta*[2] *Arne Jönsson*[2]

[1]Department of Linguistics
Göteborg University
Box 200, SE-405 30 Göteborg, Sweden
sl@ling.gu.se

[2]Department of Computer and Information Science
Linköping University
SE-581 83 Linköping, Sweden
lensa@ida.liu.se, arnjo@ida.liu.se

ABSTRACT

Distilled dialogues, i.e. re-written natural dialogues, are a useful complement to dialogues collected in Wizard of Oz-experiments or in natural settings for development of dialogue systems. However, the distillation process itself also provides insights on human-computer interaction and on properties of dialogue systems. In this paper we present the distillation process, including how the guidelines are developed, and experiences from utilising distillation.

1. INTRODUCTION

Development of dialogue systems requires knowledge on how potential users would like to interact with the system. Such knowledge is often acquired from analyses of corpora of various kinds. The two most common types of corpora are natural dialogues, collected in real settings where humans communicate in a situation similar to the one that the dialogue system is to account for, and human-computer dialogues collected in Wizard of Oz-experiments, where whole, or parts, of the dialogue system is simulated by a human intervenor and the users are given tasks to perform in an experimental set-up. The advantage of natural dialogues is that they reflect a real situation with real user tasks and goals, but it is not a human-computer interaction. Wizard of Oz-dialogues, on the other hand, have the advantage of being human-computer interaction, but are collected in an experimental setting with artificial tasks. Wizard of Oz-dialogues might also be influenced by the user's preconceptions on the conversational abilities of computers. As a complement to these two types of methods for collecting a corpus, we propose the use of distillation, i.e. re-writing natural dialogues to reflect the dialogue of a human-computer interaction [3].

Distillation is performed on corpora of transcribed natural dialogues. Guidelines are compiled, reflecting the behaviour of the dialogue system to be developed. The guidelines are then used to re-write the dialogues and a new corpus of distilled dialogues is produced, which can be utilised, for instance, for dialogue systems development.

The main issue when distilling dialogues is that in natural dialogues none of the interlocutors is a dialogue system. The system's task is instead performed by a human and the problem is how to anticipate the behaviour of a system, that does not exist, based on the performance of an agent with different performance characteristics.

It turns out that the process of distillation itself provides important insights on the properties of human-computer interaction as opposed to human interaction. This can, for instance, be used for educational purposes, or as a common ground for discussions between dialogue systems developers and researchers more focussed on investigating properties of human interaction. In this paper we discuss this aspect, i.e. distillation as a mean for understanding dialogue systems.

2. DISTILLING DIALOGUES

The process of distilling dialogues can be divided into two steps:

1. Guideline development. Guidelines are needed specifying how utterances are going to be changed. Two of the key issues when specifying guidelines are what kind of interaction they should describe and how that is achieved by a dialogue system.
2. Guideline application. When the guidelines are established they are used to distill the corpus. The question then arises how, exactly, the guidelines should be applied to a dialogue.

The distillation process is similar to empirical analyses in general. Thus, the guidelines are refined based on applying them to parts of the corpus before being finalised.

2.1. The Guidelines

Distillation guidelines depend on the properties and capabilities of the dialogue system to be developed. As the distillation method is supposed to be generic and applicable to various types of dialogue systems, we do not believe that it is feasible to specify one set of guidelines to account for all dialogue systems.

Instead, guidelines are to be developed anew for each new type of dialogue system for which distillation is carried out.

The guidelines are, of course, related to Gricean maxims and are also inspired by other guidelines on how to build "good" systems (See e.g. [4]) or how to post-process data collected in different settings (see e.g. DISC [2]).

First, there is a minimal set of general guidelines:

- Avoid changing user's utterances (the only legal reason for doing it is that the context has change and the utterance does not make sense anymore).
- Maintain dialogue coherence.
- Decide who are the parts in the dialogue. Turns made by or directed to non-participants are removed.

The first general guideline reflects the principle that we would like to develop dialogue systems that do not restrict the user's natural way of expressing themselves. The utterances produced by the interlocutor corresponding to the system's utterances are, however, modified quite frequently. The guidelines for modifying them are derived based on the properties of the dialogue system. We distinguish three different sets of properties, viz. ethical, functional and linguistic; separating content from form. E.g. it is a functional characteristic if the system gives feedback but the form it is done is a linguistic one.

Ethical properties are those aspects of behaviour that depends on ethical and politeness principles of the language community. Functional properties define the capabilities of the system, as we have mainly studied information seeking dialogues, these properties deal mostly with how information is retrieved and presented. Linguistic properties deal with how the system "express itself", e.g. whether it follows conversational conventions or if it speaks in a spoken language manner or a written language manner. If it keeps information focus or always presents data in the same manner, etc.

The properties that we have been working on are presented below. They are developed to account for a uni-modal natural language dialogue system[1]. For other dialogue systems other properties might be needed or the values need to be modified. Especially, if we are to account for a multimodal system we need to specify how various modalities are utilised by the system, which affects both the functional and linguistic properties.

Linguistic properties

- SYNTAX: S speaks syntactically correct, non-elliptically (i.e. using full sentences), and does not mumble or hesitate / S speaks according to a spoken language grammar
- TURN-TAKING(1): S does not interrupt the user
- TURN-TAKING(2): S gives up the turn if the user tries to take it
- TURN-TAKING(3): if the user does not take the turn when offered, S will give additional information if available, or otherwise be quiet
- FEEDBACK(1): S gives feedback when the user speaks
- FEEDBACK(2): S does not stop talking when the user gives feedback
- CONFIRMATION: S uses echo, implicit or explicit feedback when searching for confirmation
- TURN-KEEPING: S can keep the turn (e.g. while searching a database) by using turn-keeping devices (e.g. "let's see...")
- FOCUS: S presents the information in such a way that coherence with the focus of the user's utterance is maintained
- ADAPTATION: S adopts its lexical choices to the ones of the user
- REFORMULATION: If the user does not understand an utterance by S, S will reformulate

Functional properties

- RELEVANCE/QUANTITY(1): S only presents relevant information
- QUANTITY(2): S asks only for the information it needs to complete its task; no more and no less
- IMMEDIACY: S has immediate access to information and gives all the relevant information at once
- MEMORY: S does not forget information (e.g. does not ask for information already received)
- ORDERLINESS: S follows a certain order when asking questions; S does not skip back and forth between questions
- LIMITED INPUT ANALYSIS: S ignores aspects of input which it cannot analyse
- REPETITION: S does not repeat itself unless asked to
- USER MODEL: S behaves differently with different users
- MOTIVATION: S can motivate its actions if asked to
- CONVERSATIONAL STRATEGY: S has a good command of conversational co-operativeness, as feedback and confirmation strategies
- MAPPING: S is responsible for mapping the natural language representation of a database request to a machine suitable one

Ethical properties

- HONESTY/QUALITY: S does not lie and does not try to cheat the user
- POLITENESS: S is polite
- SERIOUSNESS: S is not ironic, does not have a sense of humor, and does not flirt

[1] Some properties are presented as positive assertions other as negative; this can be interpreted as our preferred values.

- VOLUNTARINESS: S does not try to persuade the user
- USER INITIATIVE: S does not take the initiative (including the turn) from the user
- NEUTRALITY: S does not express its own opinions
- DEFAULT ASSUMPTIONS: S can make default assumptions about the user, e.g. suggesting a particular alternative rather than providing the full range

2.2. Applying the guidelines

The next step is how, exactly, the guidelines should be applied to a dialogue. There are mainly two alternatives; either sequently applying the properties starting with distilling the whole dialogue according to ethical properties, then according to functional properties etc. or apply the whole set of properties to the utterances one at a time or to a dialogue segment at a time.

It is important to always control the coherence between utterances after distillation. Often it turns out that the application of one guideline causes the dialogue to be incoherent. Most commonly because an utterance was removed because it is subsumed by another which affects the rest of the dialogue.

To illustrate how the distillation guidelines are applied, consider the following sample of a human dialogue collected at a travel agency[2]:

9. J: what month are you leaving
10. P: well like the um third fourth of april /
11. some time around there / as cheap as
12. possible
13. J: right yes / i never heard that before /
14. the cheapest we have is um air france one
15. thousand eight hundred and ten / plus taxes
16. so that comes to / well you can have an
17. exact wait a second we'll do like this /
18. there's a tax in um both denmark and in
19. france so you'll get exact / then you need
20. one of those um international student cards
21. do you have that
22. P: no
23. J: you know what I'm referring to
24. P: yeah
25. J: ok let's see it's 90 crowns if you don't
26. have that
27. P: mm
28. J: let's see / um with taxes two thousand
29. sixty / copenhagen paris

There has been some previous turns and currently we are distilling the ninth utterance. We assume that this utterance need not be changed according to any of our distillation guidelines. It could well be that another set of guidelines stated, for instance, that the system always needs to explicitly acknowledge a user statement with an *OK*, and in that case we had modified the system's first utterance. The utterance on lines 10-12 is a user utterance and according to our general principle we do not change it unless they violate coherence criteria, which it does not.

Now, assume that we have decided that the system is not going to be ironic, i.e. conforming to the ethical property SERIOUSNESS. This then means that we remove the clerk's utterance on line 13. As we chose to set the value of the property SYNTAX to the first alternative, line 14 is modified to a full and correct sentence. Lines 15-19 are removed based on the functional property RELEVANCE. Lines 19-21 are reformulated based on the functional property QUANTITY(2) and the linguistic property SYNTAX.

Line 22 is a user utterance and it is not changed. However, the clerk's utterance on line 23 is removed based on the ethical property DEFAULT ASSUMPTION and the functional property QUANTITY(2). The DEFAULT ASSUMPTION in this case is that it is a traveling service for students, the original dialogue was recorded at the students travel agency at a small university town. This in turn makes the user utterance on line 24 incoherent and it is removed.

The functional property IMMEDIACY states that all relevant information is presented at once, thus we move the utterance in lines 25-26 and put it together with the rest of the price information. Line 27, although it is a user move, is removed because the context has changed. Finally, the clerk's last utterance, lines 28-29, is modified according to the linguistic property SYNTAX.

Thus, one possible distillation of this dialogue would yield the following (where system S has taken J:s place):

S: what month are you leaving
J: well like the um third fourth of april / some time around there / as cheap as possible
S: the cheapest we have is air france / it costs one thousand eight hundred and ten crowns plus taxes / this ticket requires an international student card / do you have one
J: no
S: then the price of a trip from copehagen to paris is two thousand sixty crowns including taxes and a student card costs 90 crowns

3. EXPERIENCES FROM DISTILLATION

The main purpose of distillation is to achieve a corpus of dialogues that can be used for development of dialogue systems. We have used distillation when developing a multimodal dialogue system [3], and we are also currently using

[2]This dialogue was collected by the University of Lund as part of the SDS project. We give an English translation of the transcription done in Göteborg as part of the same project.

the method to adapt an existing dialogue system (GoDiS, [1]) to new domains based on recorded and transcribed dialogues.

However, as discussed above, the distillation process itself facilitates the understanding of dialogue systems, both while establishing the guidelines and when applying them. We found this very useful in our own work on developing dialogue systems, and therefore decided to also use dialogue distillation for educational purposes.

We have utilised distillation on a class of 15 undergraduate students taking a research-oriented course on dialogue systems. The students were instructed to select the properties that their hypothetical system would have, and were handed a segment of a transcribed human-human dialogue to distill using the guidelines. They were encouraged to comment on and criticise the methodology, and to give suggestions for improvements.

We observed the students when they were distilling dialogues. The distillation provoked them to think, reason and argue about what properties a dialogue system should have, how they relate to properties of humans as communicating agents, and how these properties can be taxonomized and used in the process of distilling dialogues. One of the main merits of the distillation method is that it forces a confrontation between abstract principles and concrete data, by applying the distillation guidelines to actual (transcribed) human-human dialogue.

Distilling dialogues also provided the students with an understanding of issues related to utilising empirical material. For instance, there is the problem of keeping the dialogue coherent when utterances are altered. Frequently, the distiller must also alter the utterances of the "non-system" human, or the dialogue would become incomprehensible. Here, one must rely to a large part on common sense.

The students discussed the method and how that would affect the distilled dialogues. For instance, one point that had not been worked on much before the course was the distillation strategy. As alternatives to the original proposal, students suggested new distillation strategies, e.g. distilling coherent chunks of dialogue one at a time, and then possibly changing the ordering of these chunks, or starting with local properties (applying to single utterances or turns) followed by more global properties (such as memory-related factors and ordering). This in turn raised issues on how various methods would affect the distilled dialogues.

The taxonomy itself was also discussed, e.g. that there were too many properties, that they sometimes were overlapping or contradictory, or that the definitions were too vague. Some students suggested alternative taxonomies.

It was also suggested that the taxonomy may be useful as a checklist when implementing dialogue systems; they provide a way to keep track of what abilities a system has and (perhaps more importantly) what abilities are lacking.

4. CONCLUSIONS

Corpora of distilled dialogues are useful complements to corpora collected in natural settings or in Wizard of Oz-experiments. However, the process of distilling dialogues also facilitates understanding of dialogue systems. First, when the guidelines are established and second, when they are applied. This is useful not only when developing dialogue system but also for educational purposes. The distillation process stimulates discussions on what the properties of dialogue systems should be.

The method as presented in this paper is most readily applied to telephone dialogues, since these are more like (uni-modal) man-machine dialogues. This is true for the current distillation principles, but it would be possible to extend the principles to cover multi-modal phenomena such as gestures. Future research issues include the further refinement of the taxonomy, as well as the application strategy.

5. ACKNOWLEDGMENTS

This work was supported by the joint Research Program for Language Technology (HSFR/NUTEK), through the SDS project (Swedish Dialogue Systems, NUTEK/HSFR Language Technology Project F1472/1997, `http://www.ida.liu.se/~nlplab/sds/`).

The first author would like to thank the students of the Dialogue Systems course held at Göteborg University in the spring of 2000, for their valuable suggestions and comments on the distillation methodology.

6. REFERENCES

1. P. Bohlin, R. Cooper, E. Engdahl, and S. Larsson. Information states and dialogue move engines. In *Proceedings of the IJCAI-99 Workshop on Knowledge and Reasoning in Practical Dialogue Systems*, 1999.
2. DISC. Guidelines for cooperative dialogue. `http://www.disc2.dk/tools/codial/guidelinetable.html`, 2000.
3. A. Jönsson and N. Dahlbäck. Distilling dialogues - a method using natural dialogue corpora for dialogue systems development. In *Proceedings of 6th Applied Natural Language Processing Conference*, pages 44–51, 2000.
4. E. Krahmer, J. Landsbergen, X. Pouteau How to Obey the 7 Commandments for Spoken Dialogue Systems. In *Proceedings of the (E)ACL workshop on Interactive Spoken Dialog Systems*, pages 82–89, 1997.