ISSN 1744-1986

Technical Report N° 2009/ 13

Towards a Computational Pragmatics for Non-Cooperative Dialogue

Brian Pluss

18 January, 2009

Department of Computing
Faculty of Mathematics, Computing and Technology
The Open University

Walton Hall, Milton Keynes, MK7 6AA
United Kingdom

http://computing.open.ac.uk

Towards a Computational Pragmatics for Non-Cooperative Dialogue

First Year Probation Report

Brian Plüss
Computing Department
The Open University, UK

Supervisors
Paul Piwek
Richard Power

July 21, 2009

Contents

1 Introduction **2**

2 Research Question **4**
2.1 Defining Non-Cooperative Dialogue 7
2.1.1 Turn-Taking 9
2.1.2 Grounding 10
2.1.3 Speech Acts 11
2.1.4 A Note on Implicatures 12
2.1.5 Degrees of Non-Cooperation 13
2.2 Why does it matter? 15
2.3 Deliverables 16

3 Literature Review **17**
3.1 Notions of Cooperation 17
3.2 Cooperation in Dialogue Modeling 22
3.3 Approaches to Non-Cooperation in Dialogue 25
3.4 Conclusions 28

4 Methodology **29**
4.1 First Phase: The Pilot Study 29
4.1.1 Results 33
4.2 Second Phase: The Iterations 33
4.2.1 Hypothesis Formulation 34
4.2.2 Model Construction 35
4.2.3 Simulation 36
4.2.4 Evaluation 37

5 Work Plan **38**
5.1 Iterative Phase and Reflection on Results 40
5.2 Thesis Writing 42
5.3 Networking and Communication 42
5.4 Contingency Planning 43

Bibliography **44**

Chapter 1

Introduction

Most work in linguistics has approached dialogue on the assumption that participants share a common goal and cooperate to achieve it by means of conversation (Grice, 1975; Clark and Schaefer, 1989). In computational linguistics this assumption is even stronger. For instance, most dialogue systems rely on the interlocutor's full cooperation to model interaction. Consider as an example the following dialogue between a user (M) and the TRAINS-95 system (S):

```
1.   M:  I need to get to Detroit, Syracuse, and Montreal.
2.   S:  Ok.  [highlights the three cities]
3.   M:  Let's send the trains from Chicago to Detroit.
4.   S:  Ok.  [draws route from Chicago to Detroit via Toledo]
         But the terminal at Toledo is delaying traffic due
         to bad weather.  Trains will take an additional five
         hours to move through there.  [highlights Toledo red]
5.   M:  That's ok.  Now let's go from Atlanta up to Syracuse.
6.   S:  Please tell me a route to use to get from Atlanta
         to Syracuse.  [highlights engine at Atlanta and city
         Syracuse]
7.   M:  Go via Charleston and Pittsburgh.
8.   S:  Ok.  [draws route via Cincinnati] But traffic through
         Cincinnati is being delayed due to construction.
9.   M:  Ok, let's go via Charleston and Buffalo instead.
10.  S:  No problem.  [draws route]
11.  M:  That's great.
12.  S:  I hope it was helpful.
13.  M:  Finally, I want to go to Montreal from Philadelphia.
14.  S:  Ok.  [draws route]
15.  M:  That's great.  I'm done.
```

(Ferguson et al., 1996, p.73)

When expressed in normative terms, cooperation is characterised either by means of a set of imperatives or by imposing constraints on what parties

are expected to do in the interaction (Walton and Krabbe, 1995; Matheson et al., 2000). Likewise, descriptive (also called empirical or naturalistic) theories of dialogue assume cooperation as a starting point for most of their contributions (Allen and Perrault, 1980).

The research described here is aimed at the other cases, at those escaping the norms. Failure to cooperate can happen for many reasons. A non-native speaker trying to engage in a complex discussion might provide contributions which are not as clear and precise as would be expected. A student not quite sure about the topic he is supposed to elaborate on in an oral examination might provide information which is not entirely truthful or relevant. Someone suffering from dementia might produce utterances which are irrelevant or uninformative for the current exchange. These examples have to do with incompetence, ignorance and irrationality, all of which lie outside the scope of our study. We will focus on situations in which non-cooperative conversational behaviour is rational, competent and well-informed.

As an example of the kind of dialogue we will study, consider the following fragment of an interview between BBC presenter Jeremy Paxman and MP elected Greorge Galloway, aired live on 6 May, 2005:

(01)	PAXMAN:	We're joined now from his count in Bethnal Green and Bow by George Galloway. Mr Galloway, are you proud of having got rid of one of the very few black women in Parliament?
(02)	GALLOWAY:	What a preposterous question. I know it's very late in the night, but wouldn't you be better starting by congratulating me for one of the most sensational election results in modern history?
(03)	PAXMAN:	Are you proud of having got rid of one of the very few black women in Parliament?
(04)	GALLOWAY:	I'm not - Jeremy - move on to your next question.
(05)	PAXMAN:	You're not answering that one?
(06)	GALLOWAY:	No because I don't believe that people get elected because of the colour of their skin. I believe people get elected because of their record and because of their policies. So move on to your next question.
(07)	PAXMAN:	Are you proud -
(08)	GALLOWAY:	Because I've got a lot of people who want to speak to me.
(09)	PAXMAN:	- You -
(10)	GALLOWAY:	If you ask that question again, I'm going, I warn you now.
(11)	PAXMAN:	Don't try and threaten me Mr Galloway, please.
(12)	GALLOWAY:	You're the one who's trying to badger me.

(*BBC News*, 2005)

This report is part of the first-year probation assessment for a full-time Ph.D. programme. It provides details about the proposed research question (Chapter 2), a review of the relevant literature (Chapter 3), the proposed research methodology (Chapter 4) and a work plan (Chapter 5).

Chapter 2

Research Question

Computational pragmatics is the study of *language in use* or *language in context* (Levinson, 1983) from a computational perspective. This involves finding computational models of phenomena occurring in language use and evaluating the accuracy of those models (Bunt and Black, 2000).

A computational model is an abstract description of a process, system or phenomenon that can be implemented as a computer program (i.e., that is computable). Computational models usually appear as computer simulations, although model and simulation (and simulator) are different representations of the object under study. Simulations are executions of the program that implements a model, with particular input settings, directed at visualising or evaluating certain aspects of the model.

This research is aimed at answering the following question:

What properties are needed in a computational model of conversational agents so that they can engage in non-cooperative as well as in cooperative dialogue, in particular in the domain of political interviews?

Let us explain the meaning of the main terms in the statement of the question and justify their use:

- *computational model of conversational agents:* as we explained above, computational models are abstract descriptions that are computable. In our case, the model will describe the participants in a conversation (i.e., the conversational agents). This includes rules of expected behaviour for dialogues in the domain, individual goals, conversational obligations, priorities associated with goals and obligations and a dialogue management component.

- *properties:* this refers to the aspects of the model that are either general to most dialogue situations or specific to the domain in which we

have focused our research[1]. It was included in the statement of the question to indicate that it is not just a model we are after, but also an analysis of its properties in terms of how specific or generalisable they are.

- *cooperative and non-cooperative dialogue:* this distinction is central to our research and will be defined in the rest of the chapter. Intuitively, it refers to whether participants do or do not behave as is expected for the type of dialogue in which they engage. The first exchange we reproduced in the Introduction is an example of cooperative dialogue, while the fragment of the Paxman-Galloway interview is an example of non-cooperative conversation.

- *political interviews:* this is the domain in which we will focus our study of non-cooperative conversational behaviour. It is intended to provide a well-defined set of scenarios, scoping the research in a way that is suitable for a Ph.D. project.

The motivation for addressing this question is to extend the state-of-the-art of computational dialogue modeling to cases in which *things don't go that well* during a conversation. We will see what is meant by this more precisely in the next section.

The notion of cooperation between the interlocutors is at the core of most of the traditional literature on pragmatics of dialogue. For instance, Grice's notion of *conversational implicature* (Grice, 1975) provides an explanation for coherence in the following example:

Example 2.1.

> A is standing by an obviously immobilized car and is approached by B; the following exchange takes place:
>
> (1) A: *I am out of petrol.*
> B: *There is a garage round the corner.*
>
> (Grice, 1975, p.51)

Nothing in B's contribution says that the garage is open at the moment and has petrol for sale. However, as A assumes that B is trying to help, he or she can conclude that B thinks it to be the case. This information about the garage is not logically entailed by B's words; it is *implicated*.

For reasoning about such cases, Grice provides a descriptive framework, starting with the *Cooperative Principle* (CP):

[1]For example, rules of expected behaviour and obligations will apply only to dialogues in the domain, while the use of priorities associated to individual goals and obligations and the ability of agents to decide whether they will discharge an obligation or behave following their goals are properties that would apply to dialogues of any type.

> *Make your conversational contribution such as is required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which you are engaged.*
>
> (Grice, 1975, p.45)

The CP is then divided into conversational maxims: a set of rational principles which Grice grouped following the Kantian categories of *Quantity*, *Quality*, *Relation* and *Manner* (or *Modality*). We will describe Gricean pragmatics more in detail in Chapter 3, but let us look at the following as an example:

> The category of QUANTITY relates to the quantity of information to be provided, and under it fall the following maxims:
>
> 1. Make your contribution as informative as is required (for the current purposes of the exchange).
> 2. Do not make your contribution more informative than is required.
>
> (Grice, 1975, p.45)

As has been noted, e.g., by Prince (1982), despite appearing as imperatives, the CP and the maxims are of most use if regarded as presumptions that speakers and listeners can exploit for conveying and inferring meanings that are not logically entailed by their utterances. For example, if we assume the speaker is trying to help, the mention of a garage in Example 2.1, implicates that it is open and selling petrol, as otherwise he or she would be violating the Maxim of Relation (*'Be Relevant'*).

Maxims can also be *exploited*, that is *flouted* with the purpose of inducing an implicature. Consider the following:

Example 2.2.

(2) A: *How did you find the restaurant last night?*
B: *It was cheap and the toilets were clean.*

B certainly knows about, e.g., the quality of the food. The omission of this information is a violation of the Maxim of Quantity, which implicates that the food in the restaurant was not good.

Finally, speakers can also *violate* for no conversational reason, or *opt out* from a maxim in breach of the CP. Grice refers to these possibilities, although he does not elaborate on them any further:

> A participant in a talk exchange may fail to fulfill a maxim in various ways, which include the following:
>
> 1. He may quietly and unostentatiously VIOLATE a maxim; if so, in some cases he will be liable to mislead.

> 2. He may OPT OUT from the operation both of the maxim and of the CP; he may say, indicate or allow it to become plain that he is unwilling to cooperate in the way the maxim requires. He may say, for example, *I cannot say more; my lips are sealed.*
>
> (Grice, 1975, p.49)

Many revisited Grice or proposed alternatives, both in general and computational linguistics. The reliance on cooperation is even more evident in the latter, as dialogue is usually modeled in the context of dialogue systems (see Chapter 3), where users and system are expected to cooperate fully in order to achieve a purpose (e.g., booking a plane ticket). Traum, in an inspiring extended abstract (Traum, 2008), brings attention to these facts and provides reasons for focusing on non-cooperation in dialogue from a computational perspective.

2.1 Defining Non-Cooperative Dialogue

Dialogue can be accounted for either from a normative perspective, stating how participants *ought to* behave, or from a descriptive one, telling how participants *do* behave (Piwek, 2006).

In a normative approach, dialogue is modeled by means of sets of rules that participants are expected to follow when they engage in conversation (Walton and Krabbe, 1995). These sets of rules are determined by, e.g., the type of exchange, the purpose of the conversation or social conventions.

In such a context, dialogue is cooperative when participants follow the rules or stay within the restrictions imposed by these norms (see the discussion on Walton and Krabbe (1995) and Reed and Long (1997) in Chapter 3).

Consistently, in a normative framework, dialogue would be non-cooperative when participants break the rules or fail to observe the constraints of the type of exchange in which they engage. This can happen at different levels. For instance, an interviewee refusing to answer a question with sufficient reasons for doing so (i.e., opting out of the Maxim of Quantity), would be non-cooperative in the sense of Grice but cooperative in terms of the type of exchange. On the other hand, a candidate for a job implicating unwillingness to discuss his previous experience, e.g., by providing information that is not directly relevant (i.e., exploiting the Maxim of Relevance), would be cooperative in the sense of Grice but non-cooperative given the nature of the conversation.

Essential to normative approaches is that they are based on idealisations of the sort of situations they deal with. This is desirable when the aim of the approach is to formalise and allow precise reasoning, especially if we intend to implement all or part of the account at some point.

In spite of that, in this chapter we attempt to characterise a class of situations that occur in everyday, natural conversation, rather than devising a set of rules which speakers have to break in order to be considered non-cooperative. This asks for a definition, supported by empirical evidence, that would make it possible to differentiate the situations under study form others in the same context. For that reason, we will abandon the normative account (for the moment) in favour of an *operational definition* of non-cooperative dialogue. More specifically, of non-cooperative dialogue in political interviews.

The choice of the domain responds to a need for focusing the research on a well-delimited set of scenarios. The wealth of interesting conversational situations that arise in broadcast political interviews is also relevant. In the English-speaking world, and in particular in the UK, journalists are well-known for their incisive approach to public servants and political candidates. At the same time, politicians are usually well trained to deliver a set of key messages every time they speak in public, while avoiding to elaborate on issues that are not favorable to their image. We limit ourselves to considering naturally occurring (i.e. non-scripted) interviews, in which the interviewee is a politician or civil servant of any rank. We will not attempt to include exchanges with more than two main interlocutors.

Heritage (1998) analyses the distinctive organization of turn-taking in news interviews (a class of which political interviews are a proper subset). He writes[2]:

> "the participants -IRs [=interviewers] and IEs [=interviewees]- exclude themselves from a wide variety of actions that they are normally free to do in the give and take of ordinary conversation. If IRs restrict themselves to asking questions, then they cannot - at least overtly - express opinions, or argue with, debate or criticize the interviewees' positions nor, conversely, agree with, support or defend them. Correspondingly, if IEs restrict themselves to answers (or responses) to questions, then they cannot ask questions (of IRs or other IEs), nor make unsolicited comments on previous remarks, initiate changes of topic, or divert the discussion into criticisms of the IR or the broadcasting organization."
>
> (Heritage, 1998, p.8)

After this passage, the author reckons that in practice these restrictions are occasionally not observed, but only as departures from the expected behaviour and often resulting in problematic and even sanctionable courses of action.

[2]Heritage's analysis uses a broad notion of turn-taking. Instead, we will limit the use of the term to refer to the system of rules which governs how interlocutors take the floor in a conversation (Sacks et al., 1974). The *actions* mentioned by Heritage above will be regarded as speech acts.

We will structure the definition of non-cooperative dialogue in political interviews following three aspects of conversation: turn-taking, grounding and speech acts. For each group we will give a list of *non-cooperative features*. The presence and number of occurrences of these features will then determine the degree of non-cooperation in a given exchange.

2.1.1 Turn-Taking

In contrast to monologue, discourse in dialogue is constructed in turns. Speakers take turns for making their contribution at adequate places and in particular ways. Originating from the field of *conversation analysis*, it was first proposed by Sacks et al. (1974) that these shifts are governed by a set of turn-taking rules.

Interlocutors in a political interview are expected to wait for each other's turn to be over before making a shift, that is, they should respect transition-relevance places. Similarly, interviewers and interviewees are expected to start and end the interview politely or at least according to social conventions.

Definition In relation to turn-taking, a dialogue in political interviews is non-cooperative if participants:

- interrupt each other;
- speak simultaneously; or
- end the exchange abruptly.

As an example of interruption and overlapped speech, consider the following fragment from one of the interviews in our corpus (described in Chapter 4):

Example 2.3.

BBC presenter Jeremy Paxman interviews former Home Secretary Michael Howard about allegations that he had exceeded the powers of his office by instructing the Head of the Prison Service Derek Lewis[3].

(16)	HOWARD:	I did not overrule Derek Lewis-
(17)	PAXMAN:	(Interrupting) Did you threaten to overrule him?
(18)	HOWARD:	I took advice on what I could or could not do-
(19)	PAXMAN:	(Overlapping) Did you threaten to overrule him?
(20)	HOWARD:	(Overlapping) -and acted scrupulously in accordance with that advice. I did not overrule Derek Lewis-
(21)	PAXMAN:	(Interrupting) Did you threaten to overrule him?

[3]See more details and a transcript in the Appendix.

2.1.2 Grounding

Grounding, or establishing a common ground, refers to the process by which participants agree on the set of things they mutually believe, know or assume (Clark and Schaefer, 1989). In conversation, grounding is divided in two phases:

> *Presentation phase:* A presents utterance u for B to consider. He does so on the assumption that, if B gives evidence e or stronger, he can believe that she understands what he means by u.
>
> *Acceptance phase:* B accepts utterance u by giving evidence e that she believes she understands what A means by u. She does so on the assumption that, once A registers that evidence, he will also believe that she understands.
>
> (Clark and Brennan, 1991, p.130)

Regarding evidence of understanding, Clark and Schaefer (1989) identify the following methods: *continued attention*, *relevant next contribution*, *acknowledgement* (e.g., nodding), *demonstration* (e.g., paraphrasing) and *display* (e.g., verbatim repetition of each other's utterances).

In political interviews grounding is relevant to establishing a shared topical agenda. A question is accepted by providing a direct answer or by stating that the question will not be responded to. Conversely, answers are accepted by asking a next relevant question, by moving on to a new topical issue or by stating that the answer does not accord with the question asked.

Definition In relation to grounding, a dialogue in political interviews is non-cooperative if at least one of the following is observed:

After receiving an answer, the interviewer neither:

- asks a next relevant question;
- moves on to the next topical issue; nor
- states explicitly that the answer was not relevant to the previous question.

After a question, the interviewee neither:

- provides an answer which is of direct relevance; nor
- states explicitly that the question will not be answered.

As an example of these failures, consider the following fragment:

Example 2.4.

Another fragment of the Paxman-Howard interview (Example 2.3).

(11)	PAXMAN:	Did you threaten to overrule him?
(12)	HOWARD:	I was not entitled to instruct Derek Lewis and I did not instruct him.
(13)	PAXMAN:	Did you threaten to overrule him?
(14)	HOWARD:	The truth of the matter is that Mr. Marriot was not suspended-
(15)	PAXMAN:	(Interrupting) Did you threaten to overrule him?
(16)	HOWARD:	I did not overrule Derek Lewis-
(17)	PAXMAN:	(Interrupting) Did you threaten to overrule him?

2.1.3 Speech Acts

Utterances in a dialogue can be regarded as actions performed by the speaker with associated *force* and *propositional content*, or speech acts (Austin, 1962). Building on this idea, Searle (1979) proposes a classification of these acts in five classes:

- *Assertives:* giving an opinion, suggesting, concluding, etc.
- *Directives:* asking a question, ordering, requesting, inviting, etc.
- *Commissives:* promising, threatening, planning, etc.
- *Expressives:* thanking, welcoming, apologizing, etc.
- *Declarations:* declaring war, sentencing, resigning, etc.

For example, the utterance "What a preposterous question!" is an assertive speech act, with the force of giving an opinion and the propositional content of a certain question being preposterous.

Going back to the comment by Heritage quoted above, participants can fail to restrict their speech acts to those classes and associated propositional content expected for the role they assumed in the interview.

Definition In relation to speech acts, a dialogue in political interviews is non-cooperative if at least one of the following is observed:

The interviewer either:

- expresses a personal opinion;
- argues, debates with or criticises the interviewee's positions on personal (subjective) grounds; or
- agrees with, supports or defends the interviewee's positions on personal (subjective) grounds.

The interviewee either:

- asks questions of the interviewer (with the exception of clarification-request questions);
- makes (unsolicited) comments that are irrelevant to the question he was asked;
- initiates changes of topic; or
- diverts the discussion into criticisms of the interviewer or the broadcasting organization.

As an example of some of these features, consider the following fragment:

Example 2.5.

BBC presenter Jeremy Paxman interviews MP George Galloway, shortly after his victory in the UK 2005 General Election[4].

(01)	PAXMAN:	We're joined now from his count in Bethnal Green and Bow by George Galloway. Mr Galloway, are you proud of having got rid of one of the very few black women in Parliament?
(02)	GALLOWAY:	What a preposterous question. I know it's very late in the night, but wouldn't you be better starting by congratulating me for one of the most sensational election results in modern history?
		...
(30)	GALLOWAY:	You are actually conducting one of the most - even by your standards - one of the most absurd interviews I have ever participated in. I have just won an election. Can you find it within yourself to recognise that fact? To recognise the fact that the people of Bethnal Green and Bow chose me this evening. Why are you insulting them?
(31)	PAXMAN:	I'm not insulting them, I'm not insulting you
(32)	GALLOWAY:	You are insulting them, they chose me just a few minutes ago. Can't you find it within yourself even to congratulate me on this victory?

2.1.4 A Note on Implicatures

The role of implicature in defining non-cooperative dialogue in political interviews is somewhat orthogonal to the aspects we covered above.

Comments, opinions and contributions can be made either explicitly (said) or implicitly (implicated). Thus, implicature plays a functional role in determining whether an exchange is non-cooperative.

[4]The interview was aired live on 6 May, 2005 and can be found in `http://www.youtube.com/watch?v=tD5tunBGmDQ` (last access June 18 2009).

In the following example, the interviewer is expressing an opinion by flouting the Maxim of Relation:

Example 2.6.

American TV political commentator Bill O'Reilly in an interview with Hermene Hartman, the editor of an African-American newspaper in Chicago, about Obama's pastor Jeremiah Wright and his connections with Nation of Islam's leader Louis Farrakhan[5].

(27)	HARTMAN:	But what - what's the emphasis? I mean, you could also - you know, it's the twist. It's the turn that's being taken. You could also look at a wonderful sermon that Dr. Wright gave and a book developed out of it, The Audacity of Hope.
(28)	O'REILLY:	But you can't - you can't do that, though.
(29)	HARTMAN:	But we're - but here's what - you can do that if you wanted to do that.
(30)	O'REILLY:	No, no, no, no. Because every despot-
(31)	HARTMAN:	You could. Here's what - but Bill-
(32)	O'REILLY:	-and I'm not calling the man a despot, but every despot in history has done some good things. Here - look-
(33)	HARTMAN:	But he's not a despot. Come on, Bill.

The observation "every despot in history has done some good things" in (32) is not directly relevant to the conversation. The interviewer draws on the positive comment "You could also look at a wonderful sermon that Dr. Wright gave" in (27) to implicate that, despite doing good things, Dr. Wright can still be a despot, therefore expressing a personal opinion.

2.1.5 Degrees of Non-Cooperation

The previous definitions are rather strict in classifying a dialogue as non-cooperative after only one occurrence of some non-cooperative feature. This supports a view in which cooperative dialogue is regarded as the ideal situation in which participants will always do what is best to preserve the functioning of the conversation. In a political interview, this would refer to cases in which questions are answered directly and in a complete way, these answers are accepted without comments by the interviewer who goes on with the next question on the agenda, interlocutors speak in turn without interruption or overlapping, etc.

We define the degree of non-cooperation in a dialogue as the ratio between the number of occurrences of non-cooperative features present in the

[5]The full transcript of the interview and the video clip can be found in `http://mediamatters.org/items/200801220012` (last access June 18 2009).

exchange and the total number of utterances[6]. Furthermore, the degree of non-cooperation could be thus computed for the whole conversation and also for each participant, by counting only occurrences of features and utterances from each party.

In our research we will focus on (fragments of) conversations with a high degree of non-cooperation.

Definiton. The degree of non-cooperation (DNC) in a dialogue is given by the following ratio:

$$DNC = \frac{NCF}{U}$$

where NCF is the number of occurrences of non-cooperative features and U is the total number of utterances.

If we use NCF_{ir} and NCF_{ie} for referring respectively to the number of occurrences of non-cooperative features in the interviewer and the interviewee, and U_{ir} and U_{ie} for their respective number of utterances, we can compute the degree of non cooperation for each participant as follows:

$$DNC_{ir} = \frac{NCF_{ir}}{U_{ir}} \qquad\qquad DNC_{ie} = \frac{NCF_{ie}}{U_{ie}}$$

In the future, we could refine this further by associating weights to non-cooperative features and using a weighted sum instead. In that way, we could consider, for instance, that an interviewee attempting a change of subject shows a higher degree of non-cooperation than, say, one interrupting.

Example 2.7. Consider the following fragment, annotated with non-cooperative features (**O**: overlap; **GF**: grounding failure; **UC**: unsolicited comment; **I**: interruption; **TC**: topic change):

(21)	PAXMAN:	Uir.1	(Overlapping) Did you threaten to overrule him?	**O**
(22)	HOWARD:	Uie.1	-Mr. Marriot was not suspended.	**GF**
(23)	PAXMAN:	Uir.2	Did you threaten to overrule him?	**GF**
(24)	HOWARD:	Uie.2	I have accounted for my decision to dismiss Derek Lewis-	
(25)	PAXMAN:	Uir.3	(Overlapping) Did you threaten to overrule him?	**O**
(26)	HOWARD:	Uie.2	(Overlapping) -in great detail before the House of Commons-	**UC**

[6]For the moment, we will consider one non-cooperative feature per utterance. This restriction will be revisited in the future.

(27)	PAXMAN:	Uir.4	I note that you're not answering the question whether you threatened to overrule him.	
(28)	HOWARD:	Uie.3	Well, the important aspect of this which it's very clear to bear in mind-	**GF**
(29)	PAXMAN:	Uir.5	(Interrupting) I'm sorry, I'm going to be frightfully rude but -	**I**
(30)	HOWARD:	Uie.4	(Interrupting) Yes, you can -	
(31)	PAXMAN:	Uir.6	(Interrupting) I'm sorry-	**O**
(32)	HOWARD:	Uie.4	(Overlapping) - you can put the question and I will give you - I will give you an answer.	**O**
(33)	PAXMAN:	Uir.7	(Overlapping) -it's a straight yes-or-no question and a straight yes-or-no answer:	
		Uir.8	did you threaten to overrule him?	
(34)	HOWARD:	Uie.5	I discussed the matter with Derek Lewis.	
		Uie.6	I gave him the benefit of my opinion.	
		Uie.7	I gave him the benefit of my opinion in strong language, but I did not instruct him because I was not, er, entitled to instruct him.	**UC**
		Uie.8	I was entitled to express my opinion and that is what I did.	**UC**
(35)	PAXMAN:	Uir.9	With respect, that is not answering the question of whether you threatened to overrule him.	
(36)	HOWARD:	Uie.9	It's dealing with the relevant point which was what I was entitled to do and what I was not entitled to do,	**TC**
		Uie.10	and I have dealt with this in detail before the House of Commons and before the select committee.	**UC**

The following table summarises non-cooperative features, utterances and the degree of non-cooperation for each participant and for the whole fragment:

	Paxman (ir)	Howard (ie)	Fragment
Interruptions (**I**)	1	0	1
Overlaps (**O**)	3	1	4
Grounding Failure (**GF**)	1	2	3
Unsolicited Comments (**UC**)	0	4	4
Topic Change (**TC**)	0	1	1
Total NCFs	5	8	13
Utterances	9	10	19
Degree of Non-Cooperation	**0.56**	**0.80**	**0.68**

2.2 Why does it matter?

The fact that dialogue is regarded as an activity that is (or should be) inherently cooperative, has deprived alternative situations of much attention.

Preliminary findings show that implicatures, for instance, can still be used, even in cases where conversation structure is extremely flawed due to non-cooperative behaviour. This is encouraging, as it suggests that we can regard (non-)cooperation in dialogue as affecting different levels within the structure of a conversation, rather than as a binary quality that would leave outliers outside any viable characterisation.

An answer to the research question proposed above would contribute in several ways:

- *For the domain*, it would provide a better understanding of what can go wrong in political interviews and what are the possible effects of non-cooperation.
- *Direct applications* of the simulation of an accurate model of non-cooperative dialogue include training, tutoring, and many others, as noted by Traum (2008).
- *In the area of dialogue systems*, a generalization of our results would allow for the development of systems that deal with non-cooperative conversational behaviour and/or that behave non-cooperatively (according to their goals), resulting in increased flexibility and robustness.
- *As for general knowledge*, this research would provide a better understanding of dialogue structure and pragmatics by looking at cases that have not been addressed before.

2.3 Deliverables

The outcome of this research will include the following resources:

- A computational model of non-cooperative dialogue for political interviews.
- A simulator according to that model.
- An evaluation method for testing the accuracy of the model based on the output of the simulator.
- An analysis of the properties of the model and methods, with particular emphasis on how they would extend or apply to other scenarios in which non-cooperative dialogue can occur.
- The dissertation with a complete description, justification and evaluation of the work carried out during the course of the Ph.D.

Most of these items will be explained in the following chapters.

Chapter 3

Literature Review

With the aim of providing a context for this research and a justification of its relevance, this chapter presents a brief review of the literature on dialogue pragmatics and computational dialogue modeling in the light of cooperation. The discussion starts by considering different notions of cooperation from the field of pragmatics. Next, we look at cooperation in existing approaches to dialogue modeling and analyse one influential model in detail, focusing on how and why it fails to account for non-cooperative dialogue as defined in the previous section. Finally, we review other approaches to modeling non-cooperative dialogue and see how they relate to our research.

3.1 Notions of Cooperation

We have introduced the approach taken by Grice (1975) to explaining dialogue pragmatics in the previous chapter. For completeness, we repeat here the Cooperative Principle (CP) and list the four categories of conversational maxim:

> Cooperative Principle:
>
> *Make your conversational contribution such as is required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which you are engaged.*
>
> Maxim of Quantity:
>
> 1. Make your contribution as informative as is required (for the current purposes of the exchange).
> 2. Do not make your contribution more informative than is required.
>
> Maxim of Quality: Try to make your contribution one that is true.
>
> 1. Do not say what you believe to be false.
> 2. Do not say that for which you lack adequate evidence.

Maxim of Relevance: Be Relevant.

Maxim of Quality: Be perspicuous.

1. Avoid obscurity of expression.
2. Avoid ambiguity.
3. Be brief.
4. Be orderly.

(Grice, 1975, p.45-46)

Grice established a relation between the maxims and the CP: if the maxims are followed, the principle will be observed. Conversely, assuming the principle holds, observance or *exploitation* of the maxims, allows the listener to work out what the speaker is really trying to say, or *implicate*. This notion is called *conversational implicature*.

Many have criticised Grice's ideas as being too vague (e.g., Kiefer (1979), Sperber and Wilson (1982) and, more recently, Frederking (1996) and Clark (1996, p.141-146)); limited by intercultural differences (e.g., Keenan (1976), later contested by Prince (1982)); or applicable only to cases in which there is a strong sense of cooperation between the interlocutors (e.g., Asher and Lascarides (2008)).

Most of these shortcomings are evident if Grice's maxims and the CP are regarded as rules or guidelines that govern the behaviour of participants in a conversation. Grice might have contributed to the misconception, as he states: "For a time, I was attracted by the idea that observance of the CP and the maxims, in a talk exchange, could be thought of as a quasi-contractual matter, with parallels outside the realm of discourse" (Grice, 1975, p.48).

At the level of the conversation in which implicatures are used (i.e., utterance-by-utterance), participant fail to cooperate by violating or opting out from the operation of the CP or the maxims. One consequence of regarding Grice's framework as normative is that, at least to our knowledge, no attention has been paid to cases in which participants overtly or covertly step out of the operation of the CP. Note that this does not have to be always in breach of expected or rational behavior. Consider, for instance, a witness under interrogation in a U.S. trial refusing to answer a question by appealing to the Fifth Amendment of the Constitution[1]. Such behaviour will be accepted in the conversational setting as established by law, but is clearly beyond the operation of the implicature mechanism.

Prince (1982) presents a convincing analysis of the social role of the maxims in each category, emphasising that they are of most use when taken

[1] "*No person shall (...) be compelled in any criminal case to be a witness against himself*" (Bill of Rights: Amendment V, United States Constitution).

as presumptions, i.e., underlying hypotheses, that speakers and listeners can exploit for conveying and inferring meanings that are not logically entailed by the utterances. This is also the approach taken by Levinson (1983) in analysing Grice (1975).

Even though the cooperative principle does not refer explicitly to shared purposes or common goals, Grice does mention them when he elaborates on the notions of exchange and engagement. He refers to "cooperative efforts; and each participant recognises in them, to some extent, a common purpose or set of purposes, or at least a mutually accepted direction" (Grice, 1975, p.45), or assumes that "each party should, for the time being, identify himself with the transitory conversational interests of the other" (Grice, 1975, p.48).

Attardo (1997) revisits Gricean pragmatics, identifying two levels of cooperation, related to *linguistic* and *non-linguistic* goals. He claims that Grice's definition of the CP exhibits a "systematic ambiguity" in this regard. Attardo supports his view by showing how some implicatures in the examples given by Grice (1975) are worked out relying on goals that must be shared beyond that particular conversational exchange. In Attardo's view, *linguistic cooperation* refers to assumptions on the speakers' behaviour in order to encode and decode intended meaning. *Non-linguistic cooperation* is related to the behaviour of the participants towards realising the goals they intended to achieve by means of the exchange. Attardo goes on by proposing an additional principle of non-linguistic cooperation, which he calls "*Perlocutionary Cooperative Principle*" (PCP), regarding Grice's CP as relevant only to the linguistic level (an interpretation of Grice's work which is debatable). He does not claim that the theory based on Grice's CP is wrong, but instead proposes this explicit distinction between linguistic and non-linguistic cooperation as an expansion. The article includes an analysis of the relation between the PCP and other "principles" such as politeness or self-interest, which in Attardo's view override the PCP. Unfortunately, his discussion on the violation of the PCP is inconclusive and limited to observing the interdependence of the CP and PCP for working out implicatures, which follows directly from the fact that Grice intended his CP to account for both linguistic and (at least at some level of the interaction) non-linguistic cooperation. Attardo is sceptic about the possibility of meaningfully violating or flouting the PCP, somewhat close to Grice's initial temptation to regard his CP as a quasi-contractual matter. From such a normative perspective, the PCP would suffer from the same sort of criticisms and shortcomings that were identified in relation with the CP.

Along the same lines, but more amenable to computational treatment, Bunt (1994) builds on the distinction between underlying (non-conversational) and conversational goals for defining *dialogue acts*. We will come back to this in the next section.

A different approach to distinguishing between goals in conversation is that followed by Walton and Krabbe (1995). In the context of natural argumentation (or informal logic), they develop a formal, normative framework, aimed at identifying situations that lead to *fallacy*[2]. In doing so, they define a typology of dialogue, as given by the initial situation (precondition) that triggered the exchange, the joint aims assumed to be shared by the participants and their individual goals (which can be at odds). Walton and Krabbe identified six main types of dialogue in their framework: persuasion, negotiation, inquiry, deliberation, information-seeking and eristic. For example, their characterization of information-seeking dialogue is as follows (Walton and Krabbe, 1995, p.66):

Information-Seeking Dialogue

Initial Situation	Personal Ignorance
Main Goal	Spreading Knowledge & Revealing Positions
Participant's Aims	Gain, Pass on, Show, or Hide Personal Knowledge
Side Benefits	Agreement, Develop Reveal Positions, Influence Onlookers, Add to Prestige, Vent Emotions
Subtypes	Expert Consultation, Didactic Dialogue, Interview, Interrogation

Dialogues restricted to one of the six types above are called *simple dialogues*. Walton and Krabbe recognise that usually in a dialogue participants reach a situation in which the preconditions for a dialogue of a different type are met. The change from one dialogue type into another is called *dialectical shift* and must be acknowledged by both parties. When the second dialogue appears as a sub-dialogue of the first (i.e. when the second dialogue ends the first one resumes from the point where it was left) the shift is called a *functional embedding*.

Walton and Krabbe approached the problem of formally modeling persuasion dialogue from a game-theoretic perspective. They limited their study to persuasion dialogue as it was directly relevant to natural argumentation. Building on an example (presumably artificial, given the absence of sources), they identified a set of strict rules for two subtypes of persuasion dialogue (permissive and rigorous) and defined a third type as the functional embedding of a rigorous persuasion dialogue within a permissive persuasion dialogue. The dialectic shifts at the transitions between both types were also modeled by a strict set of rules, defining what was a *licit* shift and what was not. The authors claim that their model is general enough to account for dialogues like the example and their thesis is that if both participants follow the rules and all shifts are licit, the argument is valid. On the other hand, failure to follow the rules or to perform licit shifts results in fallacy.

[2]In informal logic, a fallacy is an argument that appears as valid but has flaws in the reasoning from the premises to the conclusion.

The limitations of this approach are mainly related to the lack of empirical support. Walton and Krabbe accept that a formal account can only deal with an idealization of the type of dialogue it addresses. Nevertheless, they consider the attempt to provide a mathematically precise description of (a certain class of) dialogue to be worthwhile, even if it does not apply to the vast complexity of naturally occurring conversation. We agree with this view, which is especially appealing if we attempt to approach dialogue phenomena from a computational perspective, but reckon that resulting models would be greatly improved if they were supported (and inspired) by stronger empirical evidence.

An analysis of Walton and Krabbe's framework from the perspective of cooperation appears in an article by Reed and Long (1997). The authors propose a definition of cooperation in dialogue that, they claim, acts at the level of the discourse, as opposed to the utterance-by-utterance notion of cooperation in Grice's CP. In their view, participants are cooperative if they follow a common set of dialogue rules, and stay within a mutually acknowledged framework. This means that interlocutors agree on the type of dialogue (in Walton and Krabbe's sense) in which they engage; adhere to the set of rules and main goal of that type; and also respect the rules for licit functional embedding when shifting from one type of dialogue to another.

Reed and Long also define *collaboration* as a notion stronger than cooperation, in which participant's individual goals point in the same direction (i.e., are not in conflict). They go on with an analysis of Walton and Krabbe's typology in terms of cooperation and collaboration and conclude that all dialogue falling in the classification is inherently cooperative. This is consistent with Walton and Krabbe's conclusion that, for natural argumentation, failure to follow the rules they identified for persuasion dialogue results in fallacy. However, it is not clear what the consequences are of breaking the rules for other types of dialogue and Reed and Long do not address this issue in their article.

We agree with Reed and Long in that a definition of cooperation (and therefore of non-cooperation) in these terms acts at a higher level than that of Grice. Furthermore, it allows for an analysis of participant's behaviour at the level of the discourse, from a perspective that includes obligations or presumptions determined by, e.g., social setting, participant individual goals, expected benefit from the exchange (i.e., the same aspects Walton and Krabbe take into consideration for establishing their typology of dialogues).

Bunt's (1994) treatment of dialogue acts also looks at these aspects when defining how dialogue acts operate on context. He uses the term *social context* to refer to the institutional setting of the dialogue, the roles of the participants and their communicative rights and obligations at any point in the dialogue. Individual goals are also part of the account and falls under what Bunt calls *cognitive context*.

In this section, we have seen that cooperation in dialogue can be addressed from various perspectives. The notion of expected behaviour is recurrent in the literature. Determined by low-level conversational mechanisms, social convention, institutional settings or participant's roles, underlying assumptions on speakers behaviour are relevant to the resulting structure of the conversation. This is also evident in the short examples of non-cooperative behaviour we presented in Chapter 2. Consider for instance the following from Example 2.4:

(30) GALLOWAY: You are actually conducting one of the most - even by your standards - one of the most absurd interviews I have ever participated in.

and these turns from the transcript in the Appendix:

(27) PAXMAN: I note that you're not answering the question whether you threatened to overrule him.
(. . .)
(32) HOWARD: (Overlapping) - you can put the question and I will give you - I will give you an answer.

In both cases, participants appeal to underlying assumptions on each other's expected behaviour within a political interview. When Galloway classifies the current interview as absurd, he does so by comparing it with interviews closer to an standard. Likewise, when Paxman notes that Howard is not answering the question, he does so on the basis that interviewees are expected to give an answer, as in fact Howard confirms shortly after.

As Gricean maxims are of better use taken as underlying assumptions that participants can use to enlarge the amount of information conveyed by their utterances, awareness of expected behaviour determined by the type of conversation (and awareness that those expectations can be disregarded) offers an interesting perspective from which dialogue can be analysed and modeled.

3.2 Cooperation in Dialogue Modeling

As we mentioned in previous chapters, most computational models of dialogue are constructed on the assumption that participants are cooperative (in the various ways described in the previous section). These assumptions are reasonable, as such models are developed in the context of *dialogue systems*, i.e., software systems that have a dialogue component as part of the interface.

Dialogue systems are built for the purpose of providing a service to their users. In this scenario, failure to cooperate, either on the side of the

system or of the user, is against the premises on which the system is conceived and used. Examples of dialogue systems include ARTEMIS (Sadek et al., 1997), VERBMOBIL (Wahlster, 1993) and TRAINS (Allen and Schubert, 1991). ARTEMIS is a spoken information-seeking system based on principles of rationality (Cohen et al., 1990) and intentionality (Sadek, 1992; Sadek et al., 1996). VERBMOBIL is a speech-to-speech translation system that works as a mediator between two users speaking different languages, relying on dialogue acts (Bunt, 1994) and plan recognition (Jekat et al., 1995; Alexandersson et al., 1997). TRAINS is a planning system for the transportation domain. We will analyse it in more detail in the remainder of this section, as it has been influential in developing and trying new ideas for dialogue modeling. Also, theoretical aspects of its approach to dialogue management are closely related to our research.

Before looking at the model in detail, we will introduce some concepts that are relevant both to cooperation in dialogue and to dialogue modeling, as approached in TRAINS. We will focus on dialogue acts (Bunt, 1994) and discourse obligations (Traum and Allen, 1994). Other influential notions like cognitive states (Allen and Perrault, 1980; Allen, 1995), plans (Grosz and Sidner, 1990; Chu-Carroll and Carberry, 1998) or information states (Cooper and Larsson, 1999; Matheson et al., 2000), also present in the TRAINS system, will not be part of this review but might be relevant in further stages of the project.

Dialogue acts[3] (Bunt, 1994) are an extension of speech acts to include features from turn-taking, adjacency pairs and grounding. Except for adjacency pairs, these concepts were introduced in the previous chapter, as well as their relation with cooperative conversational behaviour. Adjacency pairs, like turn-taking rules, originate from the field of conversation analysis (Schegloff, 1968) and refer to conversational structures composed of two parts produced by different speakers (e.g., question-answer, greeting-greeting, offer-acceptance/rejection).

Bunt proposes dialogue acts as "functional units used by the speaker to change the context" (Bunt, 1994, p.3) and identifies in them three properties: an *utterance form*, a *communicative function* and a *semantic context*. Changes to the context are determined by the utterance form, while communicative function and semantic content relate to the force and propositional content of a speech act, respectively. From a context-changing perspective, the communicative function determines the significance of the semantic content in the new context. Bunt gives the following example:

> For instance, a dialogue act with the utterance form "*Does it rain?*", the communicative function YES/NO QUESTION and the proposition `it is raining` as semantic content, has the effect of adding the utterance

[3] Dialogue acts are closely related to *conversational moves* (Carletta et al., 1997; Power, 1979) and *conversation acts* (Traum and Hinkelman, 1992).

> *Does it rain?* to the linguistic context, and creating in the addressee (among other things) the belief that the speaker wants to know whether the proposition `it is raining` is true.
>
> (Bunt, 1994, p.4)

The notion of context proposed by Bunt is rather elaborate, with relevant factors in five categories: *linguistic*, *semantic*, *physical*, *social* and *cognitive*. Aspects in each category are further divided into *global* and *local*, with global aspects remaining constant during the conversation and local aspects changing along the way.

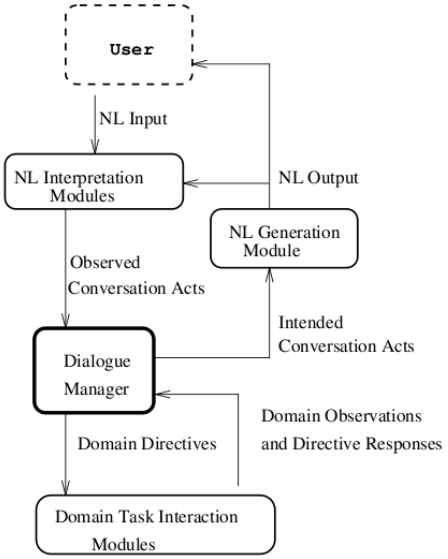
The framework is completed by distinguishing between *dialogue control acts* and *task-oriented dialogue acts*, depending on whether their communicative function is intended to control the interaction or concerned with the underlying task. Bunt analysed each of these features in detail for information-seeking dialogues, of which political interviews are a subtype. His work will certainly be relevant in the next stage of our project.

Discourse obligations (Traum and Allen, 1994) were introduced as an alternative to joint intentions (Cohen and Levesque, 1991) and shared plans (Grosz and Sidner, 1990) to allow for models of dialogue in which participants do not have the same high-level (i.e., non-conversational, individual) goals. In this view, conversational behaviour is determined, not only by participant's goals, but also by "a sense of obligation to behave within limits set by the society that the agent is part of." (Traum and Allen, 1994, p.2). Obligations are obtained from rules that encode discourse conventions and updated dynamically along the course of the conversation. In the case of conflict between goals and obligations, the latter should be favoured. The authors consider the possibility of an agent pursuing its goals at the expense of violating obligations, but they do not analyse the consequences of this any further. In fact, as we will see below, the system in which they try these ideas always discharges obligations before considering goals.

We will now analyse dialogue management aspects of the model used in the TRAINS dialogue system (Allen and Schubert, 1991) from the perspective of cooperation. The dialogue manager (see Figure 3.1) controls the structure of the conversation towards a high-level goal: obtaining a transportation plan agreed with the user according to his or her goals. The dialogue manager connects conversation acts (see footnote in the previous page) from the user with the domain tasks modules and generates the utterances for the system's next move. This is achieved by means of the algorithm in Figure 3.2, with priorities for goals, intended contributions and obligations in the following (fixed) order:

1. Discourse obligations associated with adjacency pairs
2. Weak obligation: don't interrupt the user
3. Intended contribution

4. Weak obligation: grounding
5. Discourse goals: plan negotiation
6. High-level discourse goals: form a shared plan



(Traum and Allen, 1994, p.4)

Figure 3.1: Architecture of the TRAINS Dialogue System

From the priorities and the algorithm it is clear that the system's conversational behavior is cooperative in the sense of our definition. Obligations are addressed much earlier than goals are considered. However, the approach is elegant and well suited to our question, if enriched with further types of obligations and a flexible (even dynamic) mechanism to assigning priorities to goals and obligations.

3.3 Approaches to Non-Cooperation in Dialogue

There have been previous approaches to modeling dialogue on the basis that participants are not always fully cooperative.

The first approach we will consider belongs to an area of research in which the construction of dialogue models relies strongly on a detailed description of the user, i.e., a *user model* (Wahlster and Kobsa, 1986). With the aim of analysing a user's perception of the user model in a dialogue system and the speaker's strategies for maintaining a certain image in the listener, Jameson (1989) presents an extensive study for modeling bias, individual goals, projected image and belief ascription in conversation.

```
(1)  while conversation is not finished
(2)      if system has obligations
(3)      then address obligations
(4)      else if system has turn
(5)      then if system has intended conversation acts
(6)              then call generator to produce NL utterances
(7)              else if some material is ungrounded
(8)              then address grounding situation
(9)              else if some proposal is not accepted
(10)             then consider proposals
(11)             else if high-level goals are unsatisfied
(12)             then address goals
(13)             else release turn
                      or attempt to end conversation
(14)     else if no one has turn
(15)     then take turn
(16)     else if long pause
(17)     then take turn
```

(Traum and Allen, 1994, p.5)

Figure 3.2: Discourse Actor Algorithm of the TRAINS Dialogue System

The study is realised by simulating a series of increasingly less cooperative situations in the domain of job interviews, where participants are expected to project a certain image, hide biased opinions, etc. For each utterance, the speaker selects a certain *comment*, by computing the expected *impression* it will make in the hearer, with respect to the image the speaker is trying to project. For anticipating the effect of a comment in the hearer (i.e., the impression), Jameson departs from Gricean pragmatics and proposes *pragmatic interpretations* as a basis for reasoning. He claims this mechanism to be more general than implicatures, as it does not rely on a notion of cooperation. Pragmatic interpretations operate as follows:

- Possibility p is not ruled out by comment c (possibly silence);
- but then, if p were true, comment c' would have been made instead of c, since it would have had a more desirable impact on the listener's impression;
- therefore p is apparently not realized.

Jameson implemented some of these ideas, in the context of used cars sales, by means of a dialogue system that can assume different roles (Jameson et al., 1994).

These contributions show that user-model approaches to dialogue modeling are flexible to account for situations of an arbitrary degree of intricacy. However, as noted, e.g., by Taylor et al. (1996) the level of detail required in the characterisation of the user and the complexity of mechanism for reasoning about user models can lead to problems like infinite regress in nested beliefs (speaker's beliefs about the hearer's beliefs about the speaker's beliefs...). In the same article, Taylor et al. show that nested beliefs are not necessary when participants are assumed to cooperate in the conversation, if cooperation is restricted to the absence of deception. Taylor (1994) addressed non-cooperative dialogue behaviour by implementing CYNIC, a dialogue system able to generate and recognise deception using a reasoning mechanism equivalent to a theorem prover. A notion of non-cooperation limited to deception is weaker than the one we address in this research.

More recently, Traum (2008) brought attention to the need for computational accounts of dialogue situations in which a broader notion of cooperation is not assumed. As possible applications of such models, he lists intelligent tutoring systems, bargaining agents, personal assistants acting on behalf of their owners and role-playing training agents. These applications have in common that systems can not afford to adopt the goals of their users as it happens, for instance, in the dialogue systems mentioned above. Traum also provides a list of "behaviours of interest" (along the lines of the non-cooperative features we identified in Chapter 2): unilateral topic shifts or topic maintenance, avoidance, competition, unhelpful criticism, withholding of information, lying and deception, antagonism, etc.

Traum's work on non-cooperative dialogue is mainly aimed at creating virtual humans (or embodied conversational agents (Cassell, 2001)) with abilities to engage in adversarial dialogue. Traum et al. (2005; 2008) present a model of conversation strategies for negotiation, implemented as a virtual human that can be used for teaching negotiation skills. Yet another model, with applications in the domain of training in tactical questioning, is presented by Traum et al. (2007) and Roque and Traum (2007). It can engage in dialogues like the following:

Trainee	Hello Hassan
Hassan	Hello
Trainee	How are you doing?
Hassan	Well, under the circumstances we are fine
Trainee	I'd like to talk about the marketplace
Hassan	I hope you do not expect me to tell you anything
Trainee	I just want to know why people aren't using the marketplace
Hassan	I don't feel like answering that question
Trainee	I think you know something about a tax

Hassan I am simply doing business. It is rude of you to imply otherwise

(Traum et al., 2007, p.72)

Both models include variables representing trust, politeness and emotions, and a set of strategies which are selected depending on the values of those variables. These components were developed based on studies of the respective domains and are therefore restricted to them (Roque and Traum, 2007, p.38).

It must be noted that, despite being adversarial in nature, the kind of conversational scenarios studied by Traum et al. are modeled by means of rules, i.e., the strategies for negotiation and tactical questioning identified by the authors. These rules are *followed* by the interlocutors, in accordance with the values of certain variables. This means that the dialogues accounted for by these models are adversarial but cooperative under our characterisation of non-cooperative dialogue. From this perspective, it is not clear how effective these models are to account for cases in which participants fail to follow their strategies; an issue at the core of the investigation we propose.

3.4 Conclusions

From the literature we have reviewed above, we draw the following conclusions that support and motivate our research:

- Cooperation in dialogue can be viewed from different perspectives. Depending on the type of conversational exchange in which they engage, participants assume certain bases for interaction, upon which they operate for mutual understanding. Awareness of expected behaviour and of the possibility of not meeting those expectations offers an interesting perspective from which dialogue can be analysed and modeled.

- Existing models of dialogue assume cooperation between participants, as they are developed in the context of dialogue systems, which are built with the purpose of providing a service to an interested, thus cooperative, user.

- Previous approaches to modeling dialogue without relying on full cooperation between the parties are based on rather complex descriptions of the interlocutors (user models), regard non-cooperative behaviour only as deception or account for exchanges that, yet adversarial, are cooperative in the sense that participants are following rules or strategies that describe their expected behaviour.

Chapter 4

Methodology

The overall research methodology consists of an initial exploratory study, and four subsequent stages, that can be iterated as needed to improve results. This schema is shown in Figure 4.1 and the stages described below:

1. **Exploratory Pilot Study:** aimed at scoping the research question by identifying the set phenomena to be considered in later stages.
2. **Hypothesis Formulation:** an "educated guess" explaining the phenomena observed in the pilot study and later fed into the construction of the model.
3. **Model Construction:** a computational account of the situations identified in the pilot study, which is consistent with the hypothesis.
4. **Simulation:** implementation of the model and generation of a set of dialogue situations with varying parameter settings.
5. **Evaluation:** testing the accuracy of the model (and hence the validity of the hypothesis) by comparing the outcome of the simulation with the definition given in Section 2 (internal evaluation) and by a study involving human participants (external evaluation). The results of this evaluation could in turn feed back into the hypothesis for further improvement of the model.

4.1 First Phase: The Pilot Study

Finding a computational model adequate for any sort of non-cooperative dialogue exchange was clearly beyond the possibilities of a PhD project. For this reason, we carried out a pilot study as the first stage of the investigation, principally aimed at scoping the research question.

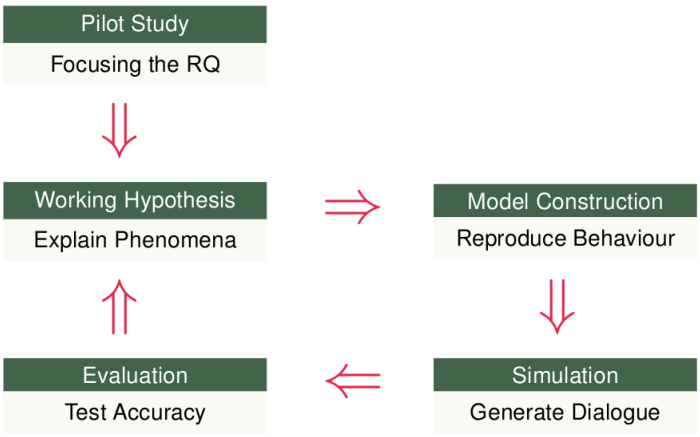


Figure 4.1: Main stages in the research methodology

These were the objectives of this phase:

- First, and most relevant, it should help focusing and narrowing the research question, by making it clear which situations will be addressed and which set of features targeted.
- Second, it should provide insight on the way non-cooperative dialogue behaviour happens and how this relates to existing computational accounts.
- Third, it should bring familiarity or proficiency with the methods and tools used for computational modeling and simulation of dialogue, and with the techniques to test the models and analyse the results.
- Fourth, it should produce an updated literature review of computational dialogue models and non-cooperative behaviour, at least for the sort of situations under study.
- Finally, criteria should emerge to allow revisiting and adapting an existing model -or to define new ones- to support non-cooperative behaviour in situations similar to those in the study.

The set of features that turn an exchange into non-cooperative dialogue were not clear at the outset of the study. They were expected to emerge from an adequate set of adversarial dialogue samples, analysed with the background provided by the literature.

For these reasons, we approached this as an empirical study, composed of the following tasks:

Task 1. Literature review.

Task 2. Preliminary data collection.

Task 3. Selection of situations and suitable samples of conflicting conversational exchange.

Task 4. Selection of one or two existing computational models of dialogue.

Task 5. Analysis of the examples for detecting features that make the exchange non-cooperative.

Task 6. Focused comparison of the samples/features with the models.

Task 7. Analysis of the results.

Task 8. Conclusions and reporting.

The plan for these tasks can be seen in Figure 4.2

The data collected for the study constitutes a corpus of 10 adversarial dialogues, totalling approximately 10,000 words. Whenever possible, the dialogues were taken from the original broadcasting media (e.g., BBC, NBC), with transcripts provided by the same source. *Youtube* and other media sharing resources were also used as preliminary sources for locating the data.

Next, the examples were analysed, identifying the features that characterised each conversation as non-cooperative (e.g., interruption, shorter turns, unfinished adjacency pairs, verbatim repetition).

At this point, looking for a better understanding of the situations under study, it was decided to preform an in-depth case study of one of the examples (see full transcript in the Appendix). A case-study approach was adequate for distinguishing the phenomena under study from others (Yin, 2003). By limiting the study to one case and considering the context in which it occurred, it was possible to approach the analysis from different angles.

By studying, e.g., the observance of turn-taking rules, the implicatures of the participants and, more extensively, how the case fitted within the normative framework proposed by Walton and Krabbe (1995), we were able to

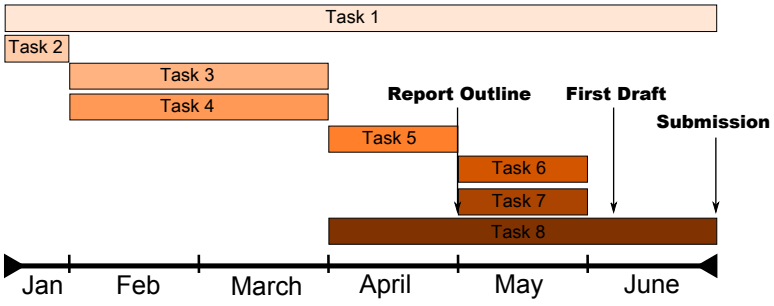


Figure 4.2: Task Schedule for the Pilot Study

better identify the nature of non-cooperative features present in the dialogue, understand the role of cooperation in the sense of Grice's Cooperative Principle and establish a formalisable framework for approaching non-cooperative dialogue towards obtaining a model. It also lead to the formulation of a preliminary hypothesis, which we expect to further substantiate and/or refine in future stages of the project.

These findings were later extended to the other cases in the corpus, increasing our confidence and ultimately leading to the methodology proposed for the second phase.

Finally, we defined non-cooperative dialogue in political interviews, taking into consideration the results from the pilot study. This definition, as given in detail in Chapter 2, was later applied to the political interviews in the corpus and also to two cases of cooperative dialogue for comparing results. The following table summarises the values obtained:

	Dialogue	Utterances	NCF	DNC
Adversarial Dialogues	1. Paxman v. Howard	54	30	**0.56**
	Paxman (IR)	24	13	**0.54**
	Howard (IE)	30	17	**0.57**
	2. Paxman v. Galloway	48	15	**0.31**
	Paxman (IR)	19	7	**0.37**
	Galloway (IE)	29	8	**0.28**
	4. O'Reilly v. Hartman	36	9	**0.25**
	O'Reilly (IR)	15	4	**0.27**
	Hartman (IE)	21	5	**0.24**
	8. Rather v. Bush (Turns 107-133)	40	19	**0.48**
	Rather (IR)	18	8	**0.44**
	Bush (IE)	22	11	**0.5**
Cooperative Dialogues	7. Keating v. Thatcher	57	0	**0**
	Keating (IR)	12	0	**0**
	Thatcher (IE)	45	0	**0**
	8. Brodie v. Blair (Turns 7-20)	31	2	**0.06**
	Brodie (IR)	9	1	**0.11**
	Blair (IE)	22	1	**0.05**

The Appendix shows the Paxman-Howard interview annotated with utterance numbers and non-cooperative features. The fragment given in Chap-

ter 2 shows how we computed the degree of non-cooperation based on the occurrence of the features.

As the table above shows, adversarial interviews present a large number of non-cooperative features, thus a high value for the degree of non-cooperation. On the other hand, cooperative exchanges have low occurrence of non-cooperative features (or none at all). These results are promising, although we reckon a need for further studies to increase our confidence in the definition and detect any shortfalls.

4.1.1 Results

The study produced the following results:

- A corpus of 10 adversarial dialogues (≈10,000 words), consisting mainly of political interviews but also of other types of exchanges (e.g., court-room interaction, entertainment interview).
- The operational definition of non-cooperative dialogue given in Chapter 2 (set of features).
- Political interviews as a domain in which to scope the research (set of scenarios).
- The review of the literature on computational dialogue models and non-cooperative conversational behaviour presented in Chapter 3.
- The methodology proposed in this chapter.
- The working hypothesis described below.

The resources of the pilot study, including the statement of objectives and provisional work plan, the corpus of adversarial dialogues, and a detailed report are available in `http://www.open.ac.uk/blogs/brianpluss/progress/pilot-study/`.

4.2 Second Phase: The Iterations

This phase follows a more conservative research approach. Given a hypothesis, we construct a model based on the hypothesis and test it by means of simulation.

As a way of avoiding the risk of working with a wrong or weak hypothesis, we will make a first iteration of the cycle considering a subset of the aspects of the domain in the model. Borrowing this rationale from early prototyping in software development, by building a partial model, implementing a simulator and evaluating the results of its output, we would quickly have a better idea on the quality of both the hypothesis and the overall approach.

The selection of aspects for this initial iteration will be based on a study as part of the substantiation of the hypothesis (see work plan in Chapter 5).

Subsequent iterations would introduce further aspects to the model. We will leave the total number of iterations undetermined, as it will depend on the results of the first iteration.

4.2.1 Hypothesis Formulation

Inspired by Walton and Krabbe's (1995) typology of dialogue and the definition of cooperation given by Reed and Long (1997), and based in the analysis of the examples in our corpus of non-cooperative dialogues, we were able to formulate a hypothesis to explain non-cooperative conversational behaviour as defined in Chapter 2.

To recap, Walton and Krabbe identify six types of dialogue (persuasion, negotiation, inquiry, deliberation, information-seeking and eristic), determined by the initial situation that triggered the exchange, the general aim of the conversation and individual goals of the participants. For each type, there is a set of strict rules interlocutors are expected to follow in order to achieve their goals by means of conversation.

By considering the following definition of cooperative dialogue (based on Reed and Long (1997)):

> A dialogue of a certain type is *cooperative* if and only if both participants follow the rules and share the joint aim of that type.

we can define non-cooperative dialogue in the following way:

> A dialogue is *non-cooperative* when (and only when) participants fail to agree on the set of rules or the joint aim of the type of dialogue in which they engage.

and formulate our hypothesis:

> *The conversational behaviour of the participants can be explained in terms of priorities associated with individual goals and conversational obligations, and of the way these priorities guide the decisions made in the course of the dialogue.*

For instance, a participant attaching high importance to individual goals might compromise the workings of a conversation by choosing contributions that go against the rules and obligations of the type of dialogue he or she is engaged in.

On the other hand, participants with higher priorities associated with the joint aims of the conversation will favour contributions consistent with the rules and obligations of the dialogue type.

It must be noted that this is a *working hypothesis*: an "educated" guess of the way participants behave in conversation in order to produce the phenomena observed in the dialogue situations defined in Section 2. The normative approach and the definition of non-cooperative dialogue given in this section are part of the hypothesis and, as such, will be subject to scrutiny, evaluation and, possibly, change for improvement in future stages of the research.

4.2.2 Model Construction

The construction of the model will be a formalization of the working hypothesis. This includes the rules for political interviews, participants' goals, conversational obligations, the priorities associated with these and a dialogue management component. The best approach for representing cognitive aspects will be the object of further investigation. Some candidates to serve as basis for our model are the BDI approach (Allen, 1995) and information states (Cooper and Larsson, 1999), which would have to be extended with the structure of priorities associated with goals and obligations, and the mechanisms for dealing with them. Discourse obligations and the model presented by Traum and Allen (1994) will certainly inform our decisions regarding the dialogue management component.

Following the insights obtained from the pilot study, at least in the first iteration, the dialogue rules for political interviews will be based on the approach taken by Walton and Krabbe (1995). In the first stages, we will model simple (one-type) dialogue, identifying intentional structures and obligations and revisiting these to account for non-cooperative behaviour. In later iterations, the model would be extended to consider different types of subdialogue, and functional embedding for, e.g., simple clarification or correction subdialogues or more complex ones like negotiation subdialogue (Chu-Carroll and Carberry, 1998).

Some attributes of the model:

- Participant will be regarded as autonomous conversational agents with a cognitive state including goals, priorities, a notion of their expected behaviour in a political interview (e.g., rules, shared goals, obligations) and some representation of the world.

- Interaction will result from independent decisions the agents take based on their cognitive state and dialogue history.

- As our focus is on local dialogue management and the resulting structure of an interaction, we will not include utterance interpretation and surface realisation as features in the model. Instead, we will use (partially) canned utterances, annotated with the information necessary to decide upon their content and force, and, possibly upon how

cooperative or not they are as alternatives for the following dialogue act.

4.2.3 Simulation

This stage involves implementing a computer program that generates dialogue according to the model constructed in the previous stage. Although the simulation will include both interlocutors, they will be autonomous in the management of the conversation and their cognitive states.

This approach to the use of simulation for the study of dialogue models was pioneered by Power (1979). It suits our project better than alternatives for various reasons. For instance, compared to Wizard-of-Oz simulations (Fraser and Gilbert, 1991), in which the behaviour described in the model is performed by a human operator, computer simulations have the following advantages:

- Computers are better at following simple rules than humans. This will be quite pressing in the first stages of the project, when conversational behaviour will be restricted to the simplest cases.
- It is easier to introduce modifications and do re-run tests. Without humans involved in the performance of the test, these can be aborted, changed and re-run as needed at a relatively low cost.
- It allows for the generation of a large number of cases, by a introducing small changes in parameter configurations (e.g., different sets of values for goal priorities). In a simulation involving humans, these subtle differences would be harder to explain or reflect in behaviour, making results less reliable.
- Computer simulators have direct applications (e.g., in trainig individuals that must put themselves in situations similar those represented in the simulation (Traum, 2008; Traum et al., 2008)).

Most of these issues apply to the decision of simulating both parties in the conversation, as opposed to taking a dialogue system approach in which humans are allowed to interact with the system.

The implementation will be carried out as a (small) software development project, following the *iterative development process* initially proposed by Basili and Turner (2005, 1975). This approach is consistent with the way we structured the research methodology and allows for early detection of any issues of concern (e.g., errors in the requirements, design or early implementation). For a model of the iterative development process, see Figure 4.3.

After each development cycle, a set of dialogues will be simulated, with different parameter configuration (e.g., different values for goal priorities).

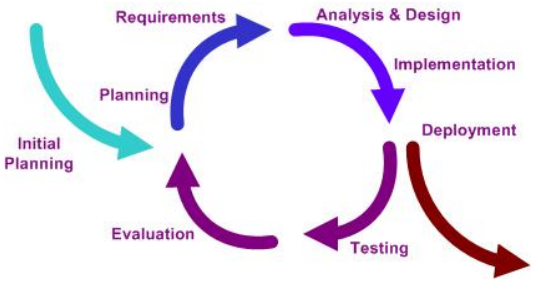


Figure 4.3: The Iterative Software Development Process

4.2.4 Evaluation

Results from the simulation in the previous stage will be evaluated internally and externally.

An internal evaluation should answer the question: *Are the dialogues generated by the simulator non-cooperative according to the definition given in Section 2?* This will involve the development of an annotation scheme following the operational definition of non-cooperative dialogue and the analysis of the output of the simulation in terms of the occurrence of non-cooperative features with respect to relevant cases in the corpus.

An external evaluation should answer the question: *Are the dialogues generated by the simulator comparable to those occurring in real conversation?* This will be performed by means of a modified Turing test. These tests are based on Turing's idea that for demonstrating intelligence, a machine's behaviour in a certain kind of interaction should not be distinguishable from that of a human (Turing, 1950). Modified Turing tests remove the interactive aspect and limit themselves to judging items produced by a computer program against similar items produced by humans. The approach is also used for evaluating results in, e.g., automated music composition (Pearce and Wiggins, 2001) and text generation (Hardcastle and Scott, 2008). If the results from a modified Turing test are inconclusive to evaluate the properties of the model[1], an alternative approach can be used. It will involve testing whether variables of the model (e.g., the degree of non-cooperation or the co-occurrence of non-cooperative features) corresponds with human judgements on the degree of cooperation of the generated dialogues.

The results from these tests will be used for further improving the hypothesis, model and simulator in subsequent iterations, until the criteria we set as completion metrics in Chapter 2 are met.

[1]This could happen, e.g., due to the use of canned utterances in the generation of the dialogues, which could draw the attention of the judges to surface attributes, rather than to non-cooperative feature in the dialogues.

Chapter 5

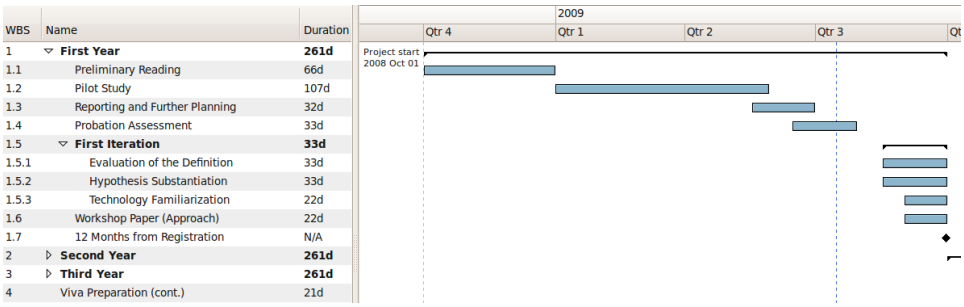
Work Plan

This chapter presents the tasks and stages planned for finding an answer to the research question stated in Chapter 2 following the methodology described in Chapter 4.

Figure 5.1 shows the plan for first year of the project, including the pilot study, the probation assessment and part of the first iteration described in Chapter 4. Provisional plans for the second and third year are shown in Figures 5.2 and 5.3, respectively.

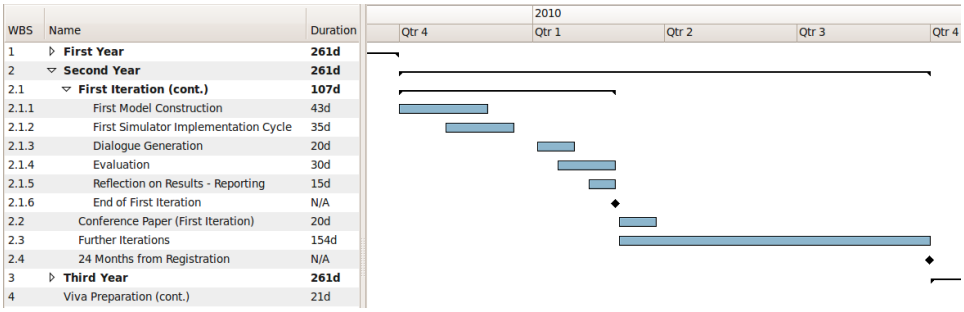
The table below shows the complete list of tasks scheduled in the plan with their estimated start date, end date and duration:

WBS	Name	Start	Finish	Duration
1	**First Year**	**Oct 1**	**Sep 30**	**261d**
1.1	Preliminary Reading	Oct 1	Dec 31	66d
1.2	Pilot Study	Jan 1	May 29	107d
1.3	Reporting and Further Planning	May 18	Jun 30	32d
1.4	Probation Assessment	Jun 15	Jul 29	33d
1.5	**First Iteration**	**Aug 17**	**Sep 30**	**33d**
1.5.1	Evaluation of the Definition	Aug 17	Sep 30	33d
1.5.2	Hypothesis Substantiation	Aug 17	Sep 30	33d
1.5.3	Technology Familiarization	Sep 1	Sep 30	22d
1.6	Workshop Paper (Approach)	Sep 1	Sep 30	22d
1.7	12 Months from Registration	Sep 30	Sep 30	N/A
2	**Second Year**	**Oct 1**	**Sep 30**	**261d**
2.1	**First Iteration (cont.)**	**Oct 1**	**Feb 26**	**107d**
2.1.1	First Model Construction	Oct 1	Nov 30	43d
2.1.2	First Simulator Implementation Cycle	Nov 2	Dec 18	35d
2.1.3	Dialogue Generation	Jan 4	Jan 29	20d
2.1.4	Evaluation	Jan 18	Feb 26	30d
2.1.5	Reflection on Results - Reporting	Feb 8	Feb 26	15d
2.1.6	End of First Iteration	Feb 26	Feb 26	N/A
2.2	Conference Paper (First Iteration)	Mar 1	Mar 26	20d
2.3	Further Iterations	Mar 1	Sep 30	154d
2.4	24 Months from Registration	Sep 30	Sep 30	N/A
3	**Third Year**	**Oct 1**	**Sep 30**	**261d**
3.1	Further Iterations (cont.)	Oct 1	Dec 23	60d
3.2	Conference Paper (Further Iterations)	Nov 15	Dec 23	29d
3.3	End of the Iterative Phase	Dec 23	Dec 23	N/A
3.4	Overall Evaluation of Methods and Results	Jan 3	Feb 28	41d
3.5	Conference Paper (Results)	Feb 1	Feb 28	20d
3.6	Analysis of Domain-Independent Properties	Feb 1	Mar 28	40d
3.7	Journal Paper (Findings and Implications)	Mar 14	May 6	40d
3.8	Journal Paper (Future Directions)	Jul 4	Aug 26	40d
3.9	Thesis Writing	Nov 1	Sep 14	228d
3.10	Submission	Sep 15	Sep 15	N/A
3.11	Viva Preparation	Sep 1	Sep 30	22d
3.12	36 Months from Registration	Sep 30	Sep 30	N/A
4	Viva Preparation (cont.)	Oct 3	Oct 31	21d



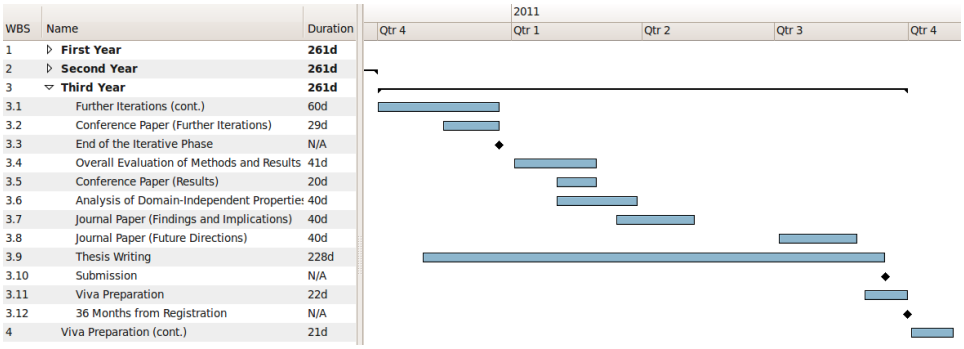
WBS	Name	Duration
1	**First Year**	**261d**
1.1	Preliminary Reading	66d
1.2	Pilot Study	107d
1.3	Reporting and Further Planning	32d
1.4	Probation Assessment	33d
1.5	**First Iteration**	**33d**
1.5.1	Evaluation of the Definition	33d
1.5.2	Hypothesis Substantiation	33d
1.5.3	Technology Familiarization	22d
1.6	Workshop Paper (Approach)	22d
1.7	12 Months from Registration	N/A
2	**Second Year**	**261d**
3	**Third Year**	**261d**
4	Viva Preparation (cont.)	21d

Figure 5.1: Work Plan for the First Year



WBS	Name	Duration
1	**First Year**	**261d**
2	**Second Year**	**261d**
2.1	**First Iteration (cont.)**	**107d**
2.1.1	First Model Construction	43d
2.1.2	First Simulator Implementation Cycle	35d
2.1.3	Dialogue Generation	20d
2.1.4	Evaluation	30d
2.1.5	Reflection on Results - Reporting	15d
2.1.6	End of First Iteration	N/A
2.2	Conference Paper (First Iteration)	20d
2.3	Further Iterations	154d
2.4	24 Months from Registration	N/A
3	**Third Year**	**261d**
4	Viva Preparation (cont.)	21d

Figure 5.2: Provisional Work Plan for the Second Year



WBS	Name	Duration
1	**First Year**	**261d**
2	**Second Year**	**261d**
3	**Third Year**	**261d**
3.1	Further Iterations (cont.)	60d
3.2	Conference Paper (Further Iterations)	29d
3.3	End of the Iterative Phase	N/A
3.4	Overall Evaluation of Methods and Results	41d
3.5	Conference Paper (Results)	20d
3.6	Analysis of Domain-Independent Properties	40d
3.7	Journal Paper (Findings and Implications)	40d
3.8	Journal Paper (Future Directions)	40d
3.9	Thesis Writing	228d
3.10	Submission	N/A
3.11	Viva Preparation	22d
3.12	36 Months from Registration	N/A
4	Viva Preparation (cont.)	21d

Figure 5.3: Provisional Work Plan for the Third Year

The following sections contain notes about these tasks, communication and networking activities and contingency planning.

5.1 Iterative Phase and Reflection on Results

As shown in Figures 5.1 and 5.2, the iterative phase spans 16.5 months, starting on the third week of August 2009 and finishing by the end of December 2010. It contains the iterations for constructing the model, implementing the simulator and evaluating the results, as described in Chapter 4. The interative phase will be followed by a complete evaluation of methods and results, from the start January to the end of February 2010, and an analysis of domain-independent properties of the model, during February and March 2010.

The first iteration was planned in detail. Those following subsequently are aggregated in Task 2.3/3.1 (Further Iterations). They will be structured in the same way, but their duration and number will be decided upon after reflecting on the results of the first iteration. This plan will be revised monthly to establish progress and make and necessary adjustments.

Without exceptions, each task will be accompanied by a review of immediately relevant literature and adequate reporting.

1.5/2.1 **First Iteration:** the first iteration will work as an early prototype of the whole process. It will start on the third week of August 2009 and finish by the end of February 2010.

1.5.1 *Evaluation of the Definition:* this task will take place only in the first iteration. It is aimed at improving understanding of political interviews, improving the non-cooperative dialogue given in Chapter 2 and testing its validity and reliability. We will design an annotation scheme for non-cooperative features and use it in a study. Subjects will be asked to annotate dialogues from our corpus and also cooperative interviews. Comparing marks from different annotators will allow us to test reliability, while correlation between non-cooperative features will serve as a test for validity. The same scheme will be used later in the evaluation task.

1.5.2 *Hypothesis Substantiation:* in the first iteration it will consist of a review of the literature relevant to the current hypothesis (e.g., rational agency and conflicting goals). In subsequent iterations it will be a revision of the hypothesis with the insight gained in the evaluation.

1.5.3 *Technology Familiarization:* a quick immersion on the technology involved in dialogue modeling (e.g., TrendiKit). This might in-

clude defining a simple model for cooperative dialogue in political interviews.

2.1.1 *First Model Construction:* a small set of dialogue rules will be devised and formalised in order to account for simple cases of interviews (e.g., question-answer sequences only).

2.1.2 *First Simulator Implementation Cycle:* the development of the simulator for the first model following the iterative development process.

2.1.3 *Dialogue Generation:* by selecting appropriate parameter settings we would simulate a set of dialogues representative of the properties of this version of the model.

2.1.4 *Evaluation:* in the first iteration it will include devising the methods for evaluating the output of the simulation, both internally against the operational definition and externally using human participants.

2.1.5 *Reflection on Results and Reporting:* an analysis of the results of the iteration and reporting by gathering the reports of each task into one document.

2.1.6 *End of First Iteration:* milestone.

2.3/3.1 **Further Iterations:** as said above, this would be a sequence of iterations spanning 10 months form March to the end of December 2010. The number and duration of each one will depend on the results of the first iteration.

3.3 **End of Iterative Phase:** milestone.

3.4 **Overall Evaluation of Methods and Results:** reflection on the results of the iterations phase and reporting.

3.6 **Analysis of Domain-Independent Properties:** identification of properties of the model that apply beyond the domain of political interviews. This would allow for an analysis of generalisable aspects of the approach and a precise description of those components that are specific to political interviews. For instance, rules of expected behaviour and obligations will certainly apply only to dialogues in the domain. On the other hand, the use of priorities associated to individual goals and obligations and the ability of agents to decide whether they will discharge an obligation or behave following their goals are properties that would apply to dialogues of any type. To support the analysis and evaluate conclusions, this stage might include a small study (one or two iterations of the methodology) in which dialogues of other types (e.g., public debates) are modeled following our approach. These dialogues would come from easily accessible resources such as the media or the internet.

5.2 Thesis Writing

Starting in parallel with the end of iterative phase, overall evaluation of the results and analysis of domain-independent properties, most of the third year will be devoted to writing the thesis and preparing for the viva.

Later on, the thesis writing task will be further refined with subtasks, milestones and deliverables, on a monthly basis. Each chapter will be dealt with in cycles with initial drafts delivered before work on the following chapter begins and revisions performed in parallel with new writing. This will allow for the detection of any inconveniences and potential delays.

5.3 Networking and Communication

We intend to expose this research to as wide a range of audiences as possible, as well as to engage in networking activities with other members of the research community.

The schedule is rather tight for the period allocated to the first iteration, which might turn publication writing and delocations difficult between August 2009 and February 2010. However, the plan is flexible for Tasks 2.3/3.1 (Further Iterations) and 3.9 (Thesis Writing) and will be able to accommodate these activities as they arise.

At present, there is one conference confirmed for September 11-12 (SIGDIAL 2009, London, UK). This is a list of other conferences that will be held within the duration of the project (some dates approximate):

- SEMDIAL: *Workshop Series on the Semantics and Pragmatics of Dialogue*. June 2010 (approximate).
- INLG 2010: *6th International Natural Language Generation Conference*. 7-9 July, 2010. Trim, County Meath, Ireland.
- ACL 2010: *The 48th Annual Meeting of the Association for Computational Linguistics*. July 11-16, 2010. Uppsala, Sweden.
- SIGDIAL 2010: *11th Meeting of the Special Interest Group on Discourse and Dialogue*. September 2010 (approximate).
- Interspeech 2010: *11th Annual Conference of the International Speech Communication Association*. 26-30 September, 2010. Makuhari, Japan.
- Interspeech 2011: *12th Annual Conference of the International Speech Communication Association*. 27-31 August, 2011. Florence, Italy.

The plan presented above includes several publications, from workshop papers to journal articles. Papers will be submitted to the conferences and workshops listed above and articles to journals relevant to the area (e.g., *Dialogue & Discourse*, *Research on Language and Computation* and the *Journal of Pragmatics*).

1.6 **Workshop Paper (Approach):** describing the problem, the approach as proposed in this report, the results of the pilot study supporting our definition of non-cooperative dialogue and the working hypothesis.

2.2 **Conference Paper (First Iteration):** with the insight obtained from the first iteration.

3.2 **Conference Paper (Further Iterations):** with the insight from the construction of the model through the iterative phase.

3.5 **Conference Paper (Results):** once the overall results and methods have been analysed.

3.7 **Journal Paper (Findings and Implications):** at the begining of the write-up task. It would be an overview of the research, describing and analysing the results and their consequences.

3.8 **Journal Paper (Future Directions):** closer to the end of the project. It would build on the results and their consequences to define future directions of work.

5.4 Contingency Planning

The best planning for contingency is the organization of the methodology, with a relatively fast initial iteration covering the whole process. This will allow for an assessment of the approach in full at an early stage and the introduction of necessary changes.

The monthly revisions in the iterative phase plan, once it be decided upon, will also contribute in detecting and correcting any problems and detours.

Bibliography

[Alexandersson et al.1997] J. Alexandersson, N. Reithinger, and E. Maier. 1997. Insights into the Dialogue Processing of VERBMOBIL. In *Proceedings of the fifth conference on Applied natural language processing*, pages 33–40. Association for Computational Linguistics Morristown, NJ, USA.

[Allen and Perrault1980] J. Allen and R. Perrault. 1980. Analyzing Intentions in Utterances. *Artificial Intelligence*, 15:143–178.

[Allen and Schubert1991] J.F. Allen and L.K. Schubert. 1991. The TRAINS project. TRAINS Technical Note 91-1. *Computer Science Dept. University of Rochester*.

[Allen1995] J. Allen. 1995. *Natural language understanding*. Benjamin-Cummings Publishing Co., Inc. Redwood City, CA, USA.

[Asher and Lascarides2008] N. Asher and A. Lascarides. 2008. Making the right commitments in dialogue. In *University of Michigan Linguistics and Philosophy Workshop, November*, volume 21-23.

[Attardo1997] S. Attardo. 1997. Locutionary and perlocutionary cooperation: The perlocutionary cooperative principle. *Journal of Pragmatics*, 27(6):753–779.

[Austin1962] J.L. Austin. 1962. *How to do things with words*. Clarendon Press.

[Basili and Turner2005] V.R. Basili and A.J. Turner. 2005. Iterative Enhancement: APractical Technique for Software Development. *Foundations of Empirical Software Engineering: The Legacy of Victor R. Basili*, page 28.

[Bunt and Black2000] H. Bunt and W. Black. 2000. *Abduction, belief and context in dialogue: studies in computational pragmatics*. John Benjamins Publishing Company.

[Bunt1994] Harry Bunt. 1994. Context and dialogue control. *THINK Quarterly*, 3.

[Carletta et al.1997] J. Carletta, S. Isard, G. Doherty-Sneddon, A. Isard, J. C. Kowtko, and A. H. Anderson. 1997. The reliability of a dialogue structure coding scheme. *Computational Linguistics*, 23(1):13–31.

[Cassell2001] J. Cassell. 2001. Embodied Conversational Agents: Representation and Intelligence in User Interfaces. *AI Magazine*, 22(4):67–84.

[Chu-Carroll and Carberry1998] J. Chu-Carroll and S. Carberry. 1998. Collaborative response generation in planning dialogues. *Computational Linguistics*, 24(3):355–400.

[Clark and Brennan1991] H.H. Clark and S.E. Brennan. 1991. Grounding in communication. *Perspectives on socially shared cognition*, pages 127–149.

[Clark and Schaefer1989] H.H. Clark and E.F. Schaefer. 1989. Contributing to discourse. *Cognitive science*, 13(2):259–294.

[Clark1996] H. H. Clark. 1996. *Using Language.* Cambridge University Press, Cambridge, MA.

[Cohen and Levesque1991] P.R. Cohen and H.J. Levesque. 1991. Confirmations and joint action. In *Proceedings of the 12 th International Joint Conference on Artificial Intelligence*, pages 951–957.

[Cohen et al.1990] P.R. Cohen, J.L. Morgan, and M.E. Pollack. 1990. *Intentions in Communication.* MIT Press.

[Cooper and Larsson1999] R. Cooper and S. Larsson. 1999. Dialogue moves and information states. *Proc. of the Third IWCS, Tilburg.*

[Ferguson et al.1996] George Ferguson, James F. Allen, and Brad Miller. 1996. Trains-95: Towards a mixed-initiative planning assistant. pages 70–77. AAAI Press.

[Fraser and Gilbert1991] N.M. Fraser and G.N. Gilbert. 1991. Simulating speech systems. *Computer Speech and Language*, 5(1):81–99.

[Frederking1996] R. Frederking. 1996. Grice's maxims: do the right thing. *Proc. of AAAI SpringSymp. on Compl. Implicature: Computational Approaches to Interpreting and Generating Conversational Implicature.*

[Grice1975] H. P. Grice. 1975. Logic and conversation. *Syntax and Semantics*, 3:41–58.

[Grosz and Sidner1990] B.J. Grosz and C.L. Sidner. 1990. Plans for discourse. *Intentions in communication*, pages 417–444.

[Hardcastle and Scott2008] D. Hardcastle and D. Scott. 2008. Can we evaluate the quality of generated text? In *The 6th edition of the Language Resources and Evaluation Conference (LREC 2008), Marrakech, Morocco.*

[Heritage1998] J. Heritage. 1998. Conversation analysis and institutional talk. Analyzing distinctive turn-taking systems. In *Proceedings of the 6 th International Congress of IADA (International Association for Dialog Analysis), Tubingen, Niemeyer.*

[Jameson et al.1994] A. Jameson, B. Kipper, A. Ndiaye, R. Schaefer, J. Simons, T. Weis, and D. Zimmermann. 1994. Cooperating to be Noncooperative: The Dialog System PRACMA. *Lecture Notes in Computer Science*, pages 106–106.

[Jameson1989] A. Jameson. 1989. But what will the listener think? Belief ascription and image maintenance in dialog. *User Models in Dialog Systems. Springer-Verlag*, pages 255–312.

[Jekat et al.1995] S. Jekat, R. Klein, E. Maier, I. Maleck, M. Mast, T. Berlin, and J.J. Quantz. 1995. Dialogue acts in VERBMOBIL.

[Keenan1976] E.O. Keenan. 1976. The universality of conversational postulates. *Language in Society*, pages 67–80.

[Kiefer1979] F. Kiefer. 1979. What do conversational maxims explain? *Linguistical Investigationes. Revue Internationale de Linguistique Française et de Linguistique Générale Paris*, 3(1):57–74.

[Levinson1983] S. C. Levinson. 1983. *Pragmatics.* Cambridge University Press.

[Matheson et al.2000] C. Matheson, M. Poesio, and D. Traum. 2000. Modelling grounding and discourse obligations using update rules. In *Proceedings of the first conference on North American chapter of the Association for Computational Linguistics*, pages 1–8. Morgan Kaufmann Publishers Inc. San Francisco, CA, USA.

[Pearce and Wiggins2001] M. Pearce and G. Wiggins. 2001. Towards a framework for the evaluation of machine compositions.

[Piwek2006] P. Piwek. 2006. Perspectives on Dialogue: Introduction to this Special Issue. *Research on Language & Computation*, 4(2):143–152, October.

[Power1979] R. Power. 1979. The organisation of purposeful dialogues. *Linguistics*, 17:107–152.

[Prince1982] Ellen F. Prince. 1982. Grice and universality: a reappraisal.

[Reed and Long1997] C. Reed and D. Long. 1997. Collaboration, cooperation and dialogue classification. *Working Notes of the IJCAI97 Workshop on Collaboration, Cooperation and Conflict in Dialogue Systems, IJCAI 97*, pages 73–78.

[Roque and Traum2007] A. Roque and D. Traum. 2007. A model of compliance and emotion for potentially adversarial dialogue agents. In *Proceedings of the 8th SIGdial Workshop on Discourse and Dialogue*.

[Sacks et al.1974] H. Sacks, E.A. Schegloff, and G. Jefferson. 1974. A simplest systematics for the organization of turn-taking for conversation. *Language*, pages 696–735.

[Sadek et al.1996] MD Sadek, A. Ferrieux, A. Cozannet, P. Bretier, F. Panaget, and J. Simonin. 1996. Effective human-computer cooperative spoken dialogue: the AGSdemonstrator. In *Spoken Language, 1996. ICSLP 96. Proceedings., Fourth International Conference on*, volume 1.

[Sadek et al.1997] M.D. Sadek, P. Bretier, and F. Panaget. 1997. ARTIMIS: Natural dialogue meets rational agency. In *international joint conference on artificial intelligence*, volume 15, pages 1030–1035. LAWRENCE ERLBAUM ASSOCIATES LTD.

[Sadek1992] M.D. Sadek. 1992. A study in the logic of intention. In *Proceedings of the 3rd Conference on Principles of Knowledge Representation and Reasoning (KR'92)*, pages 462–473.

[Schegloff1968] E.A. Schegloff. 1968. Sequencing in conversational openings. *American anthropologist*, pages 1075–1095.

[Searle1979] J.R. Searle. 1979. A Taxonomy of lllocutionary Acts. *Expression and meaning: studies in the theory of speech acts*, pages 1–29.

[Sperber and Wilson1982] D. Sperber and D. Wilson. 1982. Mutual knowledge and relevance in theories of comprehension. *Mutual knowledge*, pages 61–87.

[Taylor et al.1996] J. A. Taylor, J. Carletta, and C. Mellish. 1996. Requirements for belief models in cooperative dialogue. *User Modeling and User-Adapted Interaction*, 6(1):23–68.

[Taylor1994] JA Taylor. 1994. *A multi-agent planner for modelling dialogue*. University of Sussex.

[Traum and Allen1994] D. Traum and J. Allen. 1994. Discourse obligations in dialogue processing. In *Proceedings of the 32nd annual meeting on*

Association for Computational Linguistics, pages 1–8. Association for Computational Linguistics Morristown, NJ, USA.

[Traum and Hinkelman1992] D.R. Traum and E.A. Hinkelman. 1992. Conversation acts in task-oriented spoken dialogue. *Computational intelligence*, 8(3):575–599.

[Traum et al.2005] D. Traum, W. Swartout, S. Marsella, and J. Gratch. 2005. Fight, Flight, or Negotiate: Believable Strategies for Conversing Under Crisis. *LECTURE NOTES IN COMPUTER SCIENCE*, 3661:52.

[Traum et al.2007] D. Traum, A. Roque, A. Leuski, P. Georgiou, J. Gerten, B. Martinovski, S. Narayanan, S. Robinson, and A. Vaswani. 2007. Hassan: A virtual human for tactical questioning. In *Proceedings of the 8th SIGdial Workshop on Discourse and Dialogue, Antwerp, Belgium*, pages 71–74.

[Traum et al.2008] D. Traum, W. Swartout, J. Gratch, and S. Marsella. 2008. A virtual human dialogue model for non-team interaction. *Recent Trends in Discourse and Dialogue. Springer*.

[Traum2008] D. Traum. 2008. Extended Abstract: Computational Models of Non-cooperative dialogue. In Jonathan Ginzburg, Pat Healey, and Yo Sato, editors, *Proceedings of LONDIAL 2008, the 12th Workshop on the Semantics and Pragmatics of Dialogue*, pages 11–14, London, UK.

[Turing1950] A.M. Turing. 1950. Computing machinery and intelligence. *Mind*, 59(236):433–460.

[Wahlster and Kobsa1986] W. Wahlster and A. Kobsa. 1986. Dialogue-based user models. *Proceedings of the IEEE*, 74(7):948–960.

[Wahlster1993] W. Wahlster. 1993. *VERBMOBIL-translation of face-to-face dialogs*. Springer-Verlag London, UK.

[Walton and Krabbe1995] D. Walton and E. Krabbe. 1995. *Commitment in dialogue: Basic concepts of interpersonal reasoning*. State University of New York Press.

[Yin2003] R.K. Yin. 2003. *Applications of case study research*. Sage publications.

Appendix

The Paxman-Howard Interview

This is the transcript of an interview between BBC presenter Jeremy Paxman and former UK Home Secretary Michael Howard. The conversation refers to a meeting between Howard and the head of the Prison Service, Derek Lewis, about the dismissal of the governor of Parkhurst Prison due to repeated security failures.

The interview was first aired on 13 May 1997 and can be found at `http://www.youtube.com/watch?v=Uwlsd8RAoqI` (last access June 18 2009).

The dialogue is annotated with utterance numbers and non-cooperative features. Labels are as follows:

Uir.n**:**	interviewer's n-th utterance
Uie.n**:**	interviewee's n-th utterance
I:	interruption
O:	overlap
GF:	grounding failure
UC:	unsolicited comment
TC:	topic change

(01)	PAXMAN:	Uir.1	Right, uh . . . can you help us with this then?	
		Uir.2	You stated in your statement that the Leader of the Opposition had said that I (that is, you) personally told Mr Lewis that the governor of Parkhurst should be suspended immediately, and that when Mr Lewis objected as it was an operational matter, "I threatened to instruct him to do it".	
		Uir.3	Derek Lewis says "Howard had certainly told me that the Governor of Parkhurst should be suspended, and had threatened to overrule me".	
		Uir.4	Are you saying Mr Lewis is lying?	
(02)	HOWARD:	Uie.1	I have given a full account of this, and the position is what I told the House of Commons,	**UC**
		Uie.2	and let me tell you what the position is-	

(03)	PAXMAN:	Uir.5	(Interrupting) So you are saying that Mr Lewis lied-	**I**
(04)	HOWARD:	Uie.3	Let me tell you exactly what the position is.	
		Uie.4	I was entitled to be consulted and I was consulted,	
		Uie.5	I was entitled to express an opinion and I did express an opinion.	
		Uie.6	I was not entitled to instruct Derek Lewis what to do, and I did not instruct him what to do-	
(05)	PAXMAN:	Uir.6	(Overlapping) Well, his version-	**O**
(06)	HOWARD:	Uie.7	-and you will understand and recall that Mr Marriot was not suspended, he was moved, and Derek Lewis told the select committee of the House of Commons that it was his opinion, Derek Lewis's opinion, that he should be moved immediately.	
		Uie.8	That is what happened.	
(07)	PAXMAN:	Uir.7	Mr Lewis says "I (that is, Mr Lewis), told him what we had decided about Marriot, and why. He, (that is, you), exploded. Simply moving the governor was politically unpalatable, it sounded indecisive, it would be seen as a fudge. If I did not change my mind and suspend Marriot he would have to consider overruling me."	
(08)	HOWARD:	Uie.9	Mr Marriot-	
(09)	PAXMAN:	Uir.8	(Interrupting) You can't both be right.	**I**
(10)	HOWARD:	Uie.10	Mr Marriot was not suspended.	
		Uie.11	I was entitled to express my views,	
		Uie.12	I was entitled to be consulted-	
(11)	PAXMAN:	Uir.9	(Interrupting) Did you threaten to overrule him?	**I**
(12)	HOWARD:	Uie.13	I was not entitled to instruct Derek Lewis and I did not instruct him.	**GF**
(13)	PAXMAN:	Uir.10	Did you threaten to overrule him?	**GF**
(14)	HOWARD:	Uie.14	The truth of the matter is that Mr. Marriot was not suspended-	**GF**
(15)	PAXMAN:	Uir.11	(Interrupting) Did you threaten to overrule him?	**I**
(16)	HOWARD:	Uie.15	I did not overrule Derek Lewis-	**GF**
(17)	PAXMAN:	Uir.12	(Interrupting) Did you threaten to overrule him?	**I**
(18)	HOWARD:	Uie.16	I took advice on what I could or could not do-	**UC**
(19)	PAXMAN:	Uir.13	(Overlapping) Did you threaten to overrule him?	**O**
(20)	HOWARD:	Uie.17	(Overlapping) -and acted scrupulously in accordance with that advice.	**UC**
		Uie.18	I did not overrule Derek Lewis-	**UC**

(21)	PAXMAN:	Uir.14	(Overlapping) Did you threaten to overrule him?	**O**
(22)	HOWARD:	Uie.19	-Mr. Marriot was not suspended.	**GF**
(23)	PAXMAN:	Uir.15	Did you threaten to overrule him?	**GF**
(24)	HOWARD:	Uie.20	I have accounted for my decision to dismiss Derek Lewis-	
(25)	PAXMAN:	Uir.16	(Overlapping) Did you threaten to overrule him?	**O**
(26)	HOWARD:	Uie.20	(Overlapping) -in great detail before the House of Commons-	**UC**
(27)	PAXMAN:	Uir.17	I note that you're not answering the question whether you threatened to overrule him.	
(28)	HOWARD:	Uie.21	Well, the important aspect of this which it's very clear to bear in mind-	**GF**
(29)	PAXMAN:	Uir.18	(Interrupting) I'm sorry, I'm going to be frightfully rude but -	**I**
(30)	HOWARD:	Uie.22	(Interrupting) Yes, you can -	
(31)	PAXMAN:	Uir.19	(Interrupting) I'm sorry-	**O**
(32)	HOWARD:	Uie.22	(Overlapping) - you can put the question and I will give you - I will give you an answer.	**O**
(33)	PAXMAN:	Uir.20	(Overlapping) -it's a straight yes-or-no question and a straight yes-or-no answer:	
		Uir.21	did you threaten to overrule him?	
(34)	HOWARD:	Uie.23	I discussed the matter with Derek Lewis.	
		Uie.24	I gave him the benefit of my opinion.	
		Uie.25	I gave him the benefit of my opinion in strong language, but I did not instruct him because I was not, er, entitled to instruct him.	**UC**
		Uie.26	I was entitled to express my opinion and that is what I did.	**UC**
(35)	PAXMAN:	Uir.22	With respect, that is not answering the question of whether you threatened to overrule him.	
(36)	HOWARD:	Uie.27	It's dealing with the relevant point which was what I was entitled to do and what I was not entitled to do,	**TC**
		Uie.28	and I have dealt with this in detail before the House of Commons and before the select committee.	**UC**
(37)	PAXMAN:	Uir.23	But with respect, you haven't answered the question of whether you threatened to overrule him.	
(38)	HOWARD:	Uie.29	Well, you see, the question is what was I entitled to do and what was I not entitled to do.	**TC**
		Uie.30	I was not entitled to instruct him, and I did not do that.	**UC**
(39)	PAXMAN:	Uir.24	We'll leave that aspect there.	