Transaction and Action Coding in the Map Task Corpus*

Amy Isard and Jean Carletta
Human Communication Research Centre
University of Edinburgh
Email: J.Carletta@edinburgh.ac.uk

Abstract

Task-oriented dialogues can normally be divided into subdialogues, each of which reflects collaboration on a particular substep of the task, and which we call 'transactions'. We have devised a way of identifying transactions and their associated actions for HCRC Map Task dialogues, and we have tested the replicability of our coding scheme using naive subjects. In this report we describe the coding and the replication study.

1 Introduction

Much work on dialogue has concentrated on dialogues arising from collaborative tasks. These dialogues are both easier to analyse than free-form conversations and more relevant to practical applications of human-computer dialogue. At least from the participants' points of view, the most important issues are how to break the task into executable subtasks and what actions to perform and when. However, comparatively little work takes into account both what was said in the dialogue and the participants' actions; more often, what was said is the only evidence used. During task-oriented dialogue, the participants form collaborative plans to reach a joint goal, transferring information back and forth to each other both about how to perform the steps in the task and about the current state of the plan. Often they work through the task in an orderly fashion, performing actions as they go, and the dialogue can be divided into subsections which reflect

* This work was supported by an Interdisciplinary Research Centre Grant from the Economic and Social Research Council (U.K.) to the Universities of Edinburgh and Glasgow. The authors would like to thank Bethan Davies and Cathy Sotillo for commenting on the draft manuscript.

the steps of the task, similar to Sinclair and Coulthard's *transactions* [9]. We have devised a way of identifying transactions and their associated actions for HCRC Map Task dialogues, and we have tested the replicability of our coding scheme on naive subjects.

2 The Map Task

The HCRC Map Task Corpus [1] is a collection of 128 task-oriented dialogues involving approximately fifteen hours of speech. In the dialogues, two participants (who are either familiar or unfamiliar with each other and are separated by either a full-height or a half-height partition) have slightly different versions of a simple map with approximately fifteen landmarks on it. The maps may have different landmarks or have some of the landmarks in different locations. In addition, one participant has a route drawn on the map. The task is for the second participant to duplicate the route. An example route giver map is given in figure 1. The trials vary over whether or not the participants know each other and whether or not the partition allows eye contact. No partition allows the participants to see each other's maps.

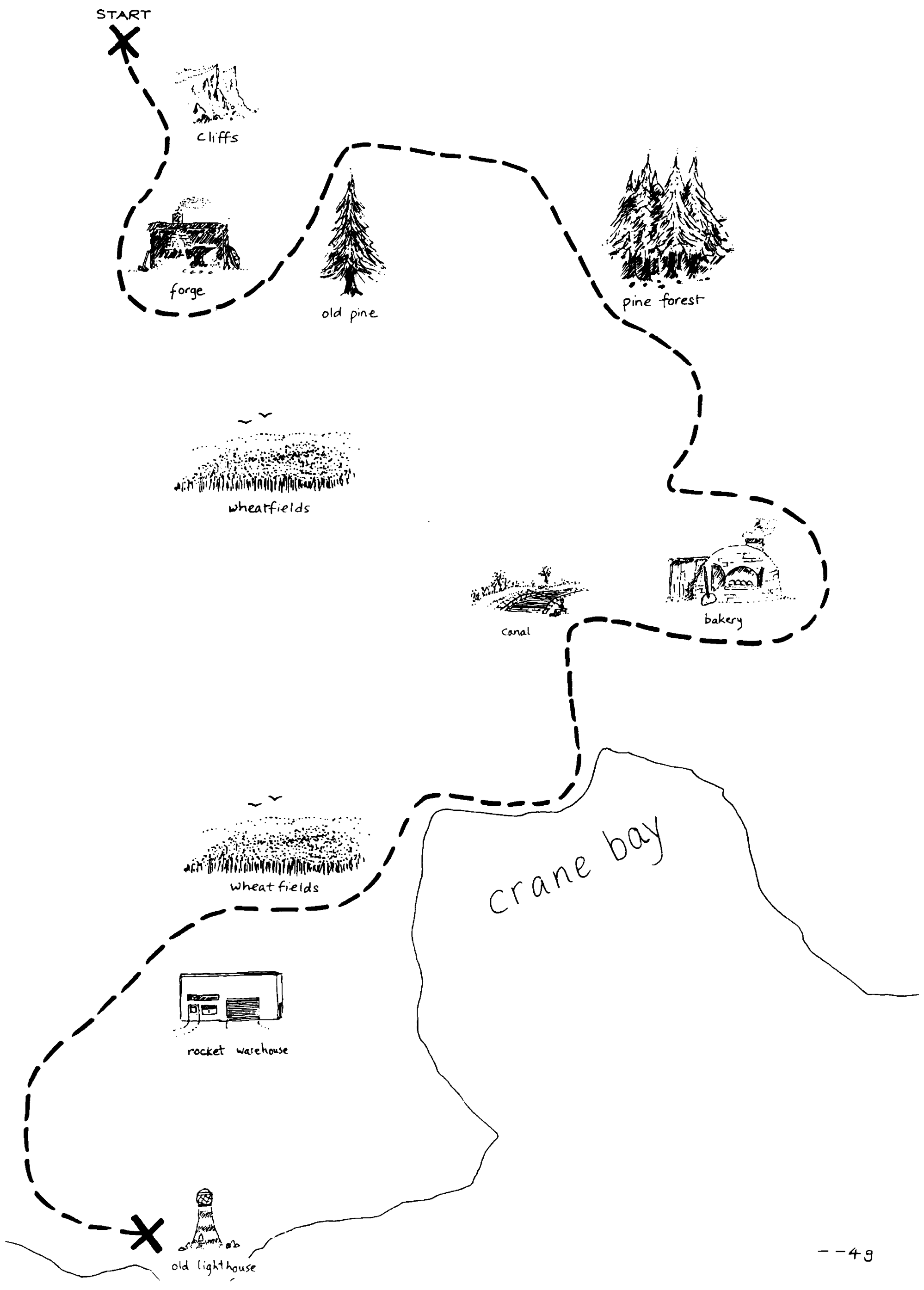
The Map Task is an ideal first domain for studying the relationship between actions and dialogue structure because there is an obvious basic order to the route which route givers almost invariably follow, even though they can choose to chunk the route in different ways. In addition, there are only two possible route follower actions (drawing a route segment and crossing a previously drawn segment out), both of which are captured on video (at least for part of the corpus).

3 The Coding System

In most map task dialogues, the route giver breaks the route into manageable segments and describes each of them one by one. Our coding system has two components: we code both (1) how route givers divide conveying the route into subtasks and what parts of the dialogue serve each of the subtasks, and (2) what actions the route follower takes and when.

Our basic route giver coding identifies the start and end of each segment and the subdialogue which conveys that route segment. However, map task participants do not always proceed along the route in an orderly fashion; as confusions arise, they often have to return to parts of the route which were previously discussed and which at least one of the participants thought had been successfully completed. In addition, participants occasionally overview an upcoming segment in order to provide a basic context for their partners, without the expectation that their partners will be able to act upon their descriptions, as in the the following transaction:

Figure 1: An Example Route Giver Map



G: And what we're basically going to be... where we're basically going to be going is towards.. I'll t-say this sort of globally

F: Mmm-hmmm.

G: then I'll do it more precisely. What we're basically doing is going, erm, south-east and then, erm, north-east, so you can imagine... a bit like a diamond shape if you like.

F: Mmm.

G: Southeast then northeast

F: Mmm.

G: and then northwest and then north, but the line's a lot more wavy than that. I'm just trying to give you some kind of overall picture.

F: Mmm.

G: It may not be very useful but.

They also sometimes engage in subdialogues which are not relevant to any segment of the route, sometimes about the experimental setup but often nothing at all to do with the task. Other types of subdialogues are possible (such as checking the placement of all map landmarks before describing any of the route, or concluding the dialogue by reviewing the entire route), but were prejudged to occur infrequently enough that we do not include them in our coding scheme. This gives us four transaction types: 'normal', 'review', 'overview', and 'irrelevant'. Coding involves marking where in the dialogue transcripts a transaction starts and which of the four types it is, and for all but 'irrelevant' transactions, indicating the start and end point of the relevant route section using numbered crosses on a copy of the route giver's map. We do not explicitly code endings of transactions because, generally speaking, transactions are large enough that they do not appear to nest; if a transaction is interrupted to, for instance, review a previous route segment, by and large participants restart the goal of the interrupted transaction afterwards rather than picking up where they left off.[1] Note that it is possible for several transactions (even of the same type) to have the same starting point on the route.

Our basic route follower coding identifies whether the follower action was drawing a segment of the route or crossing out a previously drawn segment, the start and end points of the relevant segment, indexed using numbered crosses on a copy of the route follower's map.

Appendix B explains the syntax of the coding scheme and gives an example of its use.

[1]The instructions to the naive coders allowed them to mark the ends of 'irrelevant transactions'. In practice, our expert coder never coded the endings, but always put a new transaction coding after the irrelevant section. At the most, leaving out the endings of irrelevant transactions (or any other transactions which the coder might feel are nested) loses the information about whether the following transaction is a continuation or a proper fresh start. However, because we record map reference points, we at least always know whether it is about the same segment of the route as some prior transaction.

4 The Replication Study

Since we can tell what the route follower drew and when they drew it from the video, we expect this coding to be uncontroversial. However, determining the transaction structure of a dialogue and choosing points which reflect the route segment which the route giver was trying to convey during a transaction requires judgment. In order to test whether or not our coding system for this information is replicable, we taught four naive subjects how to use it and had them all code the same four dialogues so that we could compare their coding to each other and to that of the first author. All four subjects were postgraduate students at the University of Edinburgh studying Cognitive Science, Artificial Intelligence or Maths; none of them had prior experience of the Map Task or of dialogue or discourse analysis. The four dialogues (eaq4c1-4 from the Map Task Corpus) were all from the eye contact condition but evenly mixed the other corpus conditions. All four dialogues used different maps and differently shaped routes.

We judged that the task would be too difficult if naive subjects were required to read the transcripts, listen to the tapes, and look at the maps all at the same time, so we did not give subjects access to the spoken dialogues. We prepared the dialogue transcripts so that they could be easily read, labelling each turn as coming from the giver or the follower and preserving the transcript conventions for marking overlapped speech but removing the microtagging for non-words. We did not want to allow subjects to mark transaction boundaries just anywhere in the transcripts because intonational cues can be necessary for disambiguating whether some phrases, such as 'OK' and 'right', close a transaction or open a new one. We used an existing coding of "conversational moves", described in Kowtko et al. [5], to indicate possible sites for transaction boundaries. We did not leave the move coding on the transcripts. Conversational moves divide dialogues into utterances which serve different goals in the collaboration; for instance, 'instruct' moves instruct the partner to take some action, 'query-yn' moves ask yes-no questions, and 'acknowledgement' moves show acceptance of preceding information. 'Ready' moves, a special case, are discourse markers signaling that the speaker is moving on to a new goal. It is 'ready' moves and 'acknowledgements' which are impossible to distinguish based only on written transcripts. Meanwhile, transaction boundaries can reasonably be expected to occur at move boundaries, since transactions are defined in terms of larger participant goals. Therefore we put blank lines in the transcript at each move boundary, but omitting the boundaries after ready moves. Subjects were allowed to place transaction boundaries at any of the blank lines.

Each subject was given a set of instructions with short examples (reproduced here in appendix A) and a sample dialogue extract and pair of maps to take away and read at leisure; they were asked to return with the dialogue extract coded. When they returned they were given a chance to ask questions to clarify any part of the instructions which they had not understood, and when everything had been explained to their satisfaction, they were given the four complete dialogues and maps to take away and code in their own time. The four subjects did not speak to each other about the exercise. Three of the four subjects asked for clarification of the 'overview' distinction; there were no other queries.

4.1 A note about replicability statistics

Previous work using annotative codings for discourse and dialogue has not agreed on how to show replicability of results. For instance, Kowtko et al. [5], having devised a coding scheme for conversational moves involving thirteen distinctions, cite separate pairwise agreement percentages between one expert coder and each of three naive coders in order to argue that the coding scheme is replicable. Meanwhile, Passonneau and Litman [7], in arguing that naive subjects can reliably agree on whether or not given prosodic phrase boundaries are also discourse segment boundaries, measure replicability using 'percent agreement', defined (in [4]) as the ratio of observed agreements with the majority opinion to possible agreements with the majority opinion. Neither of these approaches successfully conveys whether or not a coding scheme is replicable, most importantly because they both fail to take into account the level of agreement one would expect random coders to attain.

4.1.1 The trouble with existing measures

Chance expected agreement depends on the number and relative proportions of the categories used by the coders. For instance, suppose that two coders place objects in one of two categories in equal proportions, but choose the category randomly. We would expect them to agree with each other half of the time. If instead, the two coders were to use four categories in equal proportions, we would expect them to agree 25% of the time (since no matter what the first coder chooses, there is a 25% chance that the second coder will agree.) And if both coders were to use one of two categories, but use one of the categories 95% of the time, we would expect them to agree 90.5% of the time ($.95^2 + .05^2$ — 95% of the time the first coder chooses the first category, with a .95 chance of the second coder also choosing that category, and 5% of the time the first coder chooses the second category, with a .05 chance of the second coder also doing so). This makes it impossible to interpret raw agreement figures. It is possible to work out expected agreement for the statistic used by Passonneau and Litman but it is somewhat more difficult because the statistic itself guarantees at least 50% agreement by only pairing off coders against the majority opinion. Passonneau and Litman note that their figures are not properly interpretable and attempt to overcome this failing to some extent by showing that the agreement which they have obtained at least significantly differs from random agreement. However, all this tells us is that it is safe to assume that their coders were not coding randomly

— reassuring, but no guarantee of replicability.

4.2 The kappa statistic

Conveying the scale of agreement is inherently problematic in nominally scaled data, and no measure will allow for fully comparable results. However, our concerns are largely the same as those in the field of content analysis (see especially [6] and [10]), which has been through the same problems as we are currently facing and in which strong arguments have been made for using the kappa coefficient of agreement [8] as a measure of replicability.[2]

The kappa coefficient (K) measures pairwise agreement among a set of coders placing things into categories, correcting for expected chance agreement.

$$K = \frac{P(A) - P(E)}{1 - P(E)}$$

where $P(A)$ is the proportion of times that the coders agree and $P(E)$ is the proportion of times that we would expect them to agree by chance, calculated along the lines of the intuitive argument presented above. (For complete instructions on how to calculate K, see [8].) When there is no agreement other than that which would be expected by chance K is zero. When there is total agreement, K is one. When it is useful to do so, it is possible to test whether or not K is significantly different from chance (all of the results we cite are), but more importantly, loose interpretation of the scale of agreement is possible. Krippendorff [6] discusses what makes an acceptable level of agreement, while giving the caveat that it depends entirely on what one intends to do with the coding. For instance, he claims that it is often impossible to find associations between two variables which both rely on coding schemes with $K < .7$, and says that content analysis researchers generally think of $K > .8$ as good replicability, with $.67 < K < .8$ allowing tentative conclusions to be drawn. In addition to his caveat, we would add that coding discourse and dialogue phenomena is inherently more difficult than many previous types of content analysis (for instance, coding occurrences of particular words in a document, or the number of times that someone uses his or her telephone). Whether this makes the method inappropriate for some discourse and dialogue work remains to be seen; our point here is merely that if as a community we adopt clearer statistics, that will illuminate both our individual results and the way forward.

[2] There are several variants of the kappa coefficient in the literature, and in fact, Krippendorff's α and Siegel and Castellan's K differ slightly in the assumptions under which expected agreement is calculated. Here we use Siegel and Castellan's K, but α is so closely related in scale, especially under the usual expectations for replicability studies, that Krippendorff's statements about α hold, and we conflate the two under the more general name 'kappa'.

Figure 2: Distribution and number of transaction boundaries coded

coder	normal	review	overview	irrelevant	not a boundary
expert	47	18	0	1	591
1 (A)	46	18	10	1	582
2 (L)	52	15	3	0	587
3 (S)	47	10	1	0	599
4 (B)	27	5	2	2	621

4.3 Uses of the kappa statistic

There are several possible uses of the kappa coefficient within empirical work on discourse and dialogue, depending on whether or not one of the codings is considered definitive (the 'expert', or in some cases a coding done by a better but more expensive method). If there is no definitive coding, one might wish to describe the level of agreement amongst the coders as a group. If there is an expert coding, then it is possible to compare the group of naive coders to the expert coder both individually (in the style of Kowtko) and as a set (in the style of Passonneau and Litman, if one accepts their argument that the majority opinion actually reflects something). Comparing the naive coders to the expert as a set gives a general measure of replicability, and comparing them individually allows us to see if there are any 'odd men out' — some coders will simply do a better job than others, for whatever reason. In addition, even if there is an expert coding, it can be useful to look at agreement among the naive coders as a group. If the expert wrote the instructions, and if the naive coders agree among themselves better than they agree with the expert, it could be that the instructions mean something, but not what the expert intended! In this work we quote results for several different uses of the kappa coefficient.

4.4 Agreement on whether a move boundary is a transaction boundary

In our study, we have four naive coders and one expert classifying each move boundary as either being or not being a transaction boundary. Overall, they placed 219 transaction boundaries at 117 distinct move boundaries; there were 657 move boundaries which coders were allowed to mark. The proportions of each type of transaction boundary marked by each coder are given in figure 2 and are intended to give the reader the general flavour of the replication results.

We first look at agreement with the expert for each naive coder separately. There were 657 move boundaries for coders to mark; any one coder marked roughly one-tenth of them as transaction boundaries. Pairwise agreement, expected agreement, and kappa coefficients for each pairing of a naive coder with the expert are given in figure 3. This figure shows that, for whatever reason,

Figure 3: Agreement between the expert and naive coders individually

Coder	pairwise agreement	expected agreement	K
1 (A)	.94	.81	.68
2 (L)	.91	.81	.53
3 (S)	.94	.83	.65
4 (B)	.92	.86	.43

agreement with the expert varied considerably from subject to subject. Note that the expert only interacted minimally with the subjects, and therefore the difference does not reflect differences in the training the naive coders received.

We next look at agreement among the coders as a group. Treating our expert the same as the other coders, pairwise agreement was 93%; in other words, counting over each possible pair of coders for each of the 657 move boundaries, coders agreed on whether the move boundary was a transaction boundary 93% of the time. We would expect 83% agreement by chance from coders marking the same proportions of move boundaries as transaction boundaries but deciding which ones were transaction boundaries randomly. This gives a kappa coefficient of .59.

The previous measure checked the amount of agreement among all coders. The next question to test is whether or not our instructions conveyed what the expert coder intended by asking whether our naive coders agreed with each other more than they did with our expert coder. The results are exactly the same; 93% pairwise agreement with 83% expected by chance, with a kappa coefficient of .59. Therefore we conclude that any disagreement is general rather than reflecting a difference between expert and novice coders.

4.5 Agreement on where to place transaction boundaries

Another measure of how usable our analysis is takes the expert coder's coding as definitive and asks how often naive coders (1) failed to spot a transaction boundary or (2) marked a transaction boundary which did not exist. Defining a transaction boundary as matching only if it's in exactly the same place, our naive coders spotted our expert coder's transaction boundaries on average 60% of the time; if we omit coder 4, who tended to mark larger transactions, this rises to 67%. 33% of the boundaries which the naive coders marked in were not included by the expert coder.

4.6 Agreement on where to place points on maps

All transaction boundaries except 'irrelevant' ones have points associated with them which give the beginning and ending of the route segment under discussion. Where the coders agreed that a boundary existed, they almost invariably placed the begin and end points of their segments within the same four centimeter segment of the route, as measured on the original A3 (296 x 420 mm) maps. Many of the points were closer together than that. In contrast, the closest points which did not refer to the same begin or end point of a transaction were almost invariably more than five centimeters apart and often much further. Therefore we are confident that these points really are intended to be the same, but we can not rely on placing these points more precisely than roughly to the nearest five centimeters.

4.7 Agreement on subcodes

We believe there to be generally good agreement on which type of transaction to mark at transaction boundaries, but given the small size of this study and the fact that the vast majority of transactions are of the 'normal' type, we can not show agreement figures. There were fifteen transaction boundaries agreed by all coders; one coder marked one of them as an 'overview', but all remaining agreed boundaries were marked as 'normal' by all coders. There were 78 boundaries marked by at least two coders; of these, 64 had no dissension on the sort of boundary it was for the coders who marked it.

4.8 Sorting out disagreement

We have identified the following five areas of disagreement among our coders.

- Route givers often begin a route segment by considering possible descriptions and asking the route follower whether or not they have the landmarks involved. We intended these introductory questions to be included in the same transaction as the description which they facilitate. Sometimes coders placed the boundary after the introductory question, missing the connection with the following description. For example, in the following extract we intended a transaction boundary to occur at (1), but some coders placed it at (2).

  GIVER:

  No, it's a tiny bit.

  Okay, and while you're doing that, go in to your right a bit, but ... it doesn't really matter.

  (1)

  Okay, ehm, have you got a ravine?

  FOLLOWER:

  Yeah.

  (2)

GIVER:

Right, take that same line, still down, vertically
to the left of the ravine.

We suspect that if we were to specifically point out this situation, coders would be aware that they need to think carefully about it and they would code it more consistently. The vast majority of the disagreements of this type came from one naive coder in one dialogue, but instances of it occurred for all coders.

- Sometimes when they think a segment of the route has been finished, route followers ask questions such as "where do I go now?" We intended these questions to begin transactions, but sometimes the naive coders placed the question in the preceding transaction. For example, in the following extract we intended the boundary to be placed at (1), but some coders placed it at (2).

  FOLLOWER:

  Okay,

  (1)

  and where do we go after that?

  (2)

  GIVER:

  You turn straight up the page, and ... make a
  straight line which curves over and around the
  top of the disused monastery.

  The confusion was spread equally among the four naive coders but occurred only in one dialogue, probably reflecting that these questions are an idiosyncratic behaviour. Again, we suspect that if we were to specifically point out examples of this kind, coders would be able to mark these boundaries consistently. However, it is also possible that such questions will always be a problem area for this kind of coding, since they can simultaneously open a new transaction and close the previous one, if the route follower gives no other signal that the previous transaction's goal has been completed.

- The 'overview' coding proved to cause a lot of confusion. Subjects could not distinguish between the sequences overview-normal and normal-review where both transactions in the pair were about the same route segment. For instance, in the following extract, our expert coder marked (1) as beginning a normal transaction and (2) as beginning a review, but coder 1 marked (1) as beginning an overview and (2) as beginning a normal transaction.

GIVER:

(1)

And then, ehm, once you ... then go round the top, eh, go to your right, eh, horizontally. Go wh ... where are you at the moment, are you just at springboks here?

FOLLOWER:

Well, where I think that is, and the highest viewpoint's above where I am just now.

GIVER:

Right, okay.

(2)

Right, go vertically up then, between the springboks and highest viewpoint.

FOLLOWER:

Okay, am I meant to be trying to go under highest viewpoint? Like under

GIVER:

No no. You're meant to be going between, eh, springboks and highest p ... viewpoint.

FOLLOWER:

Okay.

We believe that the confusion stems from the fact that true overview transactions are rare, so that coders try to apply the category to more prevalent behaviours. Therefore we have decided to omit this subcode in any future coding.

- Despite our placement of possible transaction boundaries to avoid confusion about discourse markers, there were still some 'right' and 'okay' moves which had possible boundaries either side, and subjects were not able to judge where the transaction boundary belonged.[3] For instance, in the following extract, (1) and (2) might either be transaction boundaries, but not both:

GIVER:

To the left, up th ... to the right, and then curves over to the left.

(1)

FOLLOWER:

Right,

[3]There were only two such cases in this study, but we suspect that the confusion is prevalent enough overall to warrant consideration.

(2)

so we're going from like above the allotments
to just above the yacht club. No, just above ...
about, ehm, about

GIVER:

Approximately three inches,

Correcting this confusion may be a matter of giving the subjects access to the speech and pointing out that there is a distinction, if that does not make coding so difficult that the coders make more errors overall. However, many possible uses of the coding would not require boundary placement to be this precise.

- One coder had many fewer transactions than the other coders, with each transaction covering a segment of the route which other coders split into two or more transactions, indicating that he thought the route givers were planning ahead much further than the other coders did. We believe this to be a standard problem encountered by anyone involved in discourse or dialogue segmentation to which there is no fully satisfactory solution. We believe that we could improve our own situation by having a longer training period with discussion between the expert and the novice, but any such discussion would inherently make it more difficult for others to apply the coding systems and hence would weaken its credibility.

5 Conclusions

We have described a method of coding transactions and actions for the map task corpus and shown replicability results for this coding scheme using the kappa coefficient of agreement, arguing (since it is not yet well established within the dialogue community) that this is the most appropriate metric under these circumstances. Although our current level of replicability might be acceptable for some purposes, it is probably too low for establishing correlations, and this work has identified a number of confusions which it would be useful to correct at this stage. Our next step is to improve the coding scheme in an attempt to eliminate these confusions and increase its replicability.

References

[1] A. H. Anderson, M. Bader, E. G. Bard, E. Boyle, G. Doherty, S. Garrod, S. Isard, J. Kowtko, J. McAllister, J. Miller, C. Sotillo, H. Thompson, and R. Weinert. The HCRC Map Task Corpus. *Language and Speech*, 34(4):351–366, 1991.

[2] A. H. Anderson and E. A. Boyle. Forms of introduction in dialogues: Their discourse contexts and communicative consequences. *Language and Cognitive Processes*, 8:in press, 1993.

[3] J. C. Carletta. *Risk-taking and Recovery in Task-Oriented Dialogue.* PhD thesis, Edinburgh University Department of Artificial Intelligence, 1992.

[4] W. Gale, K. W. Church, and D. Yarowsky. Estimating upper and lower bounds on the performance of word-sense disambiguation programs. In *Proceedings of the Thirtieth Annual Meeting of the ACL*, June 1992.

[5] J. C. Kowtko, S. D. Isard, and G. M. Doherty. Conversational games within dialogue. Technical Report HCRC/RP-31, Human Communication Research Centre, University of Edinburgh, June 1992.

[6] K. Krippendorff. *Content Analysis: An introduction to its methodology.* Sage Publications, 1980.

[7] R. J. Passonneau and D. J. Litman. Intention-based segmentation: human reliability and correlation with linguistic cues. In *Proceedings of the 31st Annual Meeting of the ACL*, pages 148–155, June 1993.

[8] S. Siegel and J. N. J. Castellan. *Nonparametric Statistics for the Behavioral Sciences.* McGraw-Hill, second edition, 1988.

[9] J. M. Sinclair and R. M. Coulthard. *Towards an Analysis of Discourse: The English used by teachers and pupils.* Oxford University Press, 1975.

[10] R. P. Weber. *Basic Content Analysis.* Sage Publications, 1985.

A Instructions for Naive Coders

In the HCRC Map Task Dialogues, two subjects are each given similar maps like the ones shown here, one of which has a route drawn on it. They can't see each others' maps, although sometimes they can see each others' faces. The instruction-giver has the map with the route on it, and has to describe the route to the instruction-follower, who draws it on his own map.
The instruction-giver usually seems to divide the route up into pieces in her head, and describe each one in turn. Our aim is to identify the beginning and end of each of these pieces and to mark which sections of the dialogue go with which pieces of the route. In order to do this successfully, we must show that people who have never worked with the Map Task before but who are given instructions will all come to similar conclusions about the start and end point of each of these sections. We will give you four maps and whole dialogues, and ask you to divide them into sections according to the instructions below.

Here is an example from the middle of a dialogue of one piece of route:

GIVER:

    Have you got a white mountain?

FOLLOWER:

    Uh-huh.

GIVER:

    Go straight up from there, ehm, slightly curving a bit round
    the white mountain.

FOLLOWER:

    Over the top of it, uh-huh?

GIVER:

    No.

    Don't go over it. Stop when you're parallel to the top of the
    white mountain.

FOLLOWER:

    Okay. Mmhmm.

If both participants start to speak at the same time, this is marked with < at the beginning of the overlapping section and > at the end, with / marking a sentence which is carried over to the next turn. For example:

GIVER:
< Okay, I want you /

FOLLOWER:
Just below there.

GIVER:
to get to the concealed hideout. >

MARKING THE MAP

When you have worked out where the start and end points of a piece are, mark them each with a cross on the map. If you are not certain exactly where to put a cross, put it in the middle of the section where you think it might be.

Sometimes the participants think that they have completed a section but discover later that something has gone wrong, and go back to try again. This will lead to them describing the same part of the route two or more times. They may not divide the route in exactly the same way the second time they go over it, so you may have to add extra crosses.

When you have identified all the crosses, number them in order from the beginning of the route to the end.

MARKING THE TRANSCRIPT

The transcripts consist of lines of text separated by completely blank lines; please write only in the blank lines, and do not make any marks in the middle of a line of text.
Each time the participants start working on a different piece of the route, write on the transcripts just before the first thing that is said as part of the discussion:

start (start point, end point)

where start point and end point are the numbers of the crosses you have marked on the map which correspond to that particular piece of route.

You can decide where sections start and end on the transcript before you mark the map, or do both as you go along, and leave the numbering until the end, when you are sure you have decided on all the points on the map.

If it's the first time that the participants are working on that part of the route, write "normal" after the numbers

    eg start (1,2) normal

If they are reviewing a part of the route that they've already talked about, write "review" after the numbers

    eg start (5,6) review

If they are discussing something which will be dealt with later, but which the instruction-follower is not meant to draw now, write "overview" after the numbers

    eg start (3,5) overview

Sometimes one of the pair will make a comment about something which is not relevant to getting around the route, on a subject like the weather, the shape of the room, or the pen they are using. In this case, mark the beginning of the section

    "start irrelevant"

and the end of it

    "end irrelevant".

If they are in the middle of a discussion of a part of the route when they talk about the irrelevant topic, and they continue this discussion afterwards as if nothing had happened, there is no need to start a new numbered segment. So you could end up with something like

start (2,3) normal

```
GIVER:
        talk talk
FOLLOWER:
        talk talk


start irrelevant

GIVER:
        irrelevant talk
FOLLOWER:
        irrelevant talk

end irrelevant

GIVER:
        talk talk
FOLLOWER
        talk talk

start (3,4) normal
```

So you will end up with the transcript and map looking something like this:

When you have finished reading the instructions, we will give you part of a dialogue to mark, so that you can come back to us with any questions before we send you away with the four whole dialogues. The same two people were not involved in all four dialogues you will be given; four people each had a turn at being instruction-giver and instruction-follower.

The whole dialogues will be numbered 1 to 4; please do them in this order.

B The Coding in the Transcript Files

Transaction and action information is given on new coding lines, indicated by the unique identifiers *A and *B.[4] Each *A line marks the beginning of a transaction and gives the coordinates of the start and end points of the route segment which the route giver is trying to convey by that transaction. Our coding scheme does not allow transactions to overlap, so putting ends in the files is redundant. Each *A line takes one of the following forms, where things which are underlined are literal:

For normal transactions: <u>*A</u> index point-one<u>,</u>point-two

For review transactions: <u>*A</u> index point-one<u>,</u>point-two <u>r</u>

For non-goal-directed segments: <u>*A</u> index <u>irrelevant</u>

Each index is a unique positive integer; the indices start at 1 at the beginning of each dialogue and increment. The points are currently represented as indices to points drawn by hand on a hardcopy of the vmaps.[5]

*B lines are placed to the nearest turn where the route follower actually begins to draw the line on the map:

For drawing: <u>*B</u> point-one<u>,</u>point-two

For scribbling out: <u>*B</u> point-one<u>,</u>point-two <u>scribble</u>

Note that because often turns overlap and because drawing can begin between turns, placement of *B lines can only be approximate. Drawing *B lines give the points where the route follower started and ended the line, indexed on a hardcopy of the route giver's map. Erasing *B lines give the points describing the line which was scribbled out at that time.

For now, we expect to have to determine which *B lines go with which *A lines by hand; this way of coding the lines does not guarantee a straightforward relationship. Later we could modify the coding so that each *B line contains the index of the *A transaction which it is associated with.

[4]Currently (6 January 1995) eaq4 and naq4 have been coded on line for *A lines. All map reference points are on hardcopies of the maps. Eaq4c1 - eaq4c4 have been *B coded on paper copies of the transcripts and maps.

[5]We could eventually represent the points as (x,y) coordinates, where x and y give the distance in centimetres from the left and bottom of the A3 page, respectively, rounded to the nearest centimetre. Before we do, we want to figure out how we're going to use the points. We need to be able to relate them to the route and deviation scores easily. Either we will always have to do it by hand, in which case we may as well leave the system as it is now, or we will have to write a graphical interface for the points and the route which can calculate deviation scores automatically.

As an example, we include eaq4c1 as coded by the first author. We apologise for poor reproduction quality of the maps and the fact that the maps are not available in electronic versions of this report.

```
*TA 1
*A 1 1,2
Start at the extinct volcano,
and go down round the tribal settlement.
And then

*TB 2
Whereabouts is the tribal settlement?

*TA 3
It's at the bottom.
It's to the left of the {a e} extinct volcano.

*TB 4
*B 1,2
Right.
How far?

*TA 5
{m Ehm}, at the opposite side.

*TB 6
To the opposite side.
Is it underneath the rope bridge
or to the {a lef}

*TA 7
It's underneath the rope bridge.
*A 2 2,3
*B 2,3
And then from the tribal settlement go straight up towards the rope
bridge
and over the rope bridge.
*B 3,4
Then down three steps
and along to above the volcano.

*TB 8
{x Eh}, {a d} ... Is down three steps below or above the machete?

*TA 9
{x Ah}. The machete's not on my map.
```

*TB 10
{x Oh}.

*TA 11
*B 4,5
Down three lines.

*TB 12
Right.

*TA 13
*B 5,6
And then along as far as the volcano
but above it,
*B 6,7
and stop underneath the collapsed shelter
but away from it a bit,

*TB 14
Right.

*TA 15
*A 3 3,4
And go up to about the middle of the map.

*TB 16
*B 7,8
The middle of the map.

*TA 17
And stop.

*TB 18
Just slightly above the crevasse?

*TA 19
That's not on my map either.
*A 4 4,5
*B 8,9
{m Ehm}, go to your left again into about the middle.

*TB 20
I think that would bring me over the crevasse.

*TA 21
Well, it's not on my map.

*TB 22
No? {x Oh}.

*TA 23
*B 9,10
Well. Go a quarter of the way along the page
and down again.

*TB 24
*B 10,11
That'll take me right into it.

*TA 25
Well.
Go right along instead then.

*TB 26
Along to the left?

*TA 27
Is the Saxon barn on your map?

*TB 28
*B 8,12
Yes.

*TA 29
Well, go along to the Saxon barn,
*A 5 5,6
and down past the pelicans.

*TB 30
*B 9,11 scribble
*C scribbles out entire T-shaped mistake involving 9,10,11
Not marked on my map.

*TA 31
{m Ehm}, go down from the Saxon barn, {m ehm},

*TB 32
< To the /

*TA 33
{a T}

*TB 34
left or right of it? >

*TA 35
Straight down to the left.

*TB 36
Left of it, right.

*TA 37
And just above the rope bridge
and then along to your left again.

*TB 38
*B 12,13
Okay.

*TA 39
*A 6 6,7
And from there go straight up to the top of the map
to the golden beach at the top left-hand corner,
passing the white mountain.

*TB 40
{u But} golden beach is in the top right-hand corner.

*TA 41
< Left-hand corner, sorry.
{x Oh}. There's /

*TB 42
But there's a

*TA 43
two golden beaches on this map as well. >

*TB 44
Right, above the white mountain?

*TA 45
*B 13,14
{g Mmhmm}.

And then stop.
*A 7 7,8
Then turn to your right,

*TB 46
Right.

*TA 47
and go along to the middle of the map.

*TB 48
{a B} ... {x Eh}, underneath slate mountain?

*TA 49
No.
Just to the {a e} ... the end of slate mountain.

*TB 50
*B 14,15
Right.

*TA 51
*A 8 8,9
*B 15,16
Then go over slate mountain,

*TB 52
Right.

*TA 53
*A 9 9,10
And turn to your left, into the middle of the map.
That'd be your right, I suppose.

*TB 54
To my right.
Now very close to the right of the page.

*TA 55
No.
Left then.

*TB 56
Then whereabouts?

*TA 57
Into the middle.
*A 10 10,11
*B 16,17
Then stop
and go straight down at the side of the secret valley.

*TB 58
Left or right side?

*TA 59
Left.

*TB 60
*B 17,18
Left side.
Right, okay.

*TA 61
*B 18,19
And go along underneath the secret valley
and finish.

*TB 62
Right.