which recipe to use to achieve a discourse purpose before the implementation of the recipe itself is discussed. So, the IU for deciding a recipe generally comes before the IU(s) on implementing the recipe. In some cases, however, the recipe itself will be built from scratch, and elaborated in parallel with implementation of the plan. In addition, the recursive step of breaking down a plan into subplans is not strictly temporally ordered with respect to choosing a recipe and implementing it. Often, once a joint purpose is established, participants may iterate through various parts of the negotiation process such as considering and rejecting several recipes for achieving the joint purpose, or may even encounter the failure of a subplan, which then necessitates replanning (see Section 5.4).

Finally, the last step in the schema, in which dialogue participants explicitly confirm their achievement of (or failure to achieve) their joint purpose, seems to frequently arise in the genre of cooperative problem-solving dialogues. These exchanges are also called pre-closings [Schegloff and Sacks, 1973]. In macro-level analysis, there is the opportunity for goal confirmation at various levels of planning. For example, in the TRAINS dialogues, participants often explicitly confirm that a substep in their joint plan meets a temporal constraint, and then later confirm that the entire plan has been solved. At the lowest level of confirmation, e.g. tokens realizing the speech act of acknowledgment, however, no macro-level coding is required since these types of utterance-token exchanges are already grouped together into a single CGU. We suggest the convention that confirmation subdialogues be embedded in the IU that concerns the most recent action(s) contributing to the outcome being evaluated. For example, in the IU analysis in (32), in iu.1.3 the partners plan to move engine E2 to Corning. In iu.1.3.1, they check that the overall time constraint can still be satisfied, if engine E2 is moved according to their plan. In this example, iu.1.3.1 is embedded relative to iu.1.3,

---

and second, the remaining labels of relevance, “Assess Situation” and “Establish Solution”, seem to conflate the distinction between decomposing a plan into subplans and choosing and implementing recipes to achieve a given plan. For example, segments in which dialogue participants propose either components to a solution (i.e. subplans) or resources to accomplish a task are both labeled as “Establish Solution” segments. Following our proposed schema, the former would be part of step 3 in the process, “Negotiate how to achieve joint goal”, while the latter would be part of step 2 in that same process. The recursion in the proposed schema falls out naturally from this distinction. In this regard, the schema maps more closely to the *discourse script* for the TRAINS domain used in [Traum, 1994], since both explicitly address goal decomposition.

since the time constraint checked in iu.1.3.1 is evaluated with respect to the action recently planned out in iu.1.3.

It can be seen from the schema presented that the informational-intentional interface relies heavily on the distinct notions of plan (a mental state to intend to achieve a goal) versus recipe (a data structure description for action to achieve a specified goal) [Pollack, 1986, Bratman, 1987]. When dialogue participants are explicitly discussing their goals or subgoals, the IU takes on a more intentional flavor. When participants are identifying parameters for executing recipes, the IU takes on a more informational flavor.

We will have to discover whether the application of macro-level analysis to additional genres of corpora, such as conversational and tutoring dialogues, can be handled by extensions to this schema or will require the formulation of entirely new schemata.

5.4 Replanning and common ground corrections

Analogous to *cancelled* CGUs at the meso-level, we posit the notion of *modified* IUs at the macro-level. A modified IU is defined as any IU whose grounded content is replaced or corrected, in part or in whole, by the content introduced into common ground by later IUs. These later IUs will be referred to as *correcting* or *modifying* IUs. Modifying IUs may typically concern either changes in beliefs or changes in goals or plans. For example, a change in belief may be expressed when a mistake in reference is discovered, such as confusing engine E1 with engine E2. Or it may occur when a genuine discrepancy in understanding arises, as when an automatic spoken dialogue agent (or human!) misrecognizes *Newark* as *New York* and proceeds to provide the wrong itinerary information. A change in goal may occur when dialogue participants discover a plan constraint cannot be satisfied or enablement conditions do not hold and cannot be brought about.

For the sake of efficiency, only the most specific IUs that are modified by a correcting IU will be marked by placing the symbol `#` before the IU heading. In certain cases, multiple IUs must be marked as modified, as in (34).

(34) iu.1 plan weekend activity
    #iu.1.1 plan to drive in A's car
    #iu.1.2 plan to picnic at lakeside
  .
  .
  .
  iu.5 plan to go to seashore instead of lake
    iu.5.1 plan to take train because A's car will be in car shop

In other cases, an IU whose content is only partially modified by later correcting IUs will be marked, as in (35).

(35) #iu.1 plan to drive in A's car to lakehouse for the weekend
  .
  .
  iu.5 plan to meet to catch the train instead
    iu.5.1 plan to take train because A's car will be in car shop

In general, if multiple IUs at lower-levels of the IU tree contain all occurrences of modified content, mark the multiple lower IUs individually. More general IUs whose content is only partially modified should be labeled with the symbol # only when no other marking of lower-level IUs containing the equivalent content is possible.

5.5 Comparative notes

As noted in the Introduction to this coding manual, there are many aspects of discourse structure that remain unanalyzed at the macro-level. Issues of individual agency and interaction management are but two areas that remain unaddressed. While we envision many possible extensions to this guide to capture additional dimensions of macro-structure, we would also like to highlight what is new in our approach toward informational and intentional analysis.

The use of CGUs as the basic units for organizing the macro-level structure is not significant for the form of the analysis. From the point of view of interpreting the content, however, using CGUs instead of the utterance-tokens themselves actually solves some outstanding problems with the use of strictly tree-based linguistic structures that organize textual units themselves, as is prevalent in current methods of macro-level analysis.

The first issue that is solved is that of temporal discontinuities of utterance-tokens that share or continue the same topic or purpose. The temporally unrestricted assignment of utterance-tokens to CGU "sets" allows for the recording of such discontinuities, yet makes them invisible at the macro-level. These discontinuities, if important to understand, can be retrieved from the meso-level CGU coding and reasoned about. Discourse phenomena involving discontinuities include discourse interruptions [Grosz and Sidner, 1985], or "out-of-order" and "out-of-the-blue" segments [Nakatani *et al.*, 1995], and even crossing dependencies which arise not infrequently when parallel topics are interleaved in discussion (See e.g. CGUs 21 and 22 (24)), encoding the Trains dialogue given in Appendix A.2.

A second significant issue addressed by using CGUs rather than utterance-tokens as the basic units is that of ambiguous utterance-tokens. As shown in the example meso-level analyses, it is quite common to have an utterance-token belong to more than one CGU (e.g. when it affirms understanding of a question and then adds new information). These meaningful ambiguities, or multiple discourse roles, of certain kinds of utterance-tokens are one source of the problem of "fuzzy boundaries" in text-based discourse segmentation [Passonneau and Litman, 1997]. Another source of fuzzy boundaries is the mandatory assignment to discourse segments of rather meaningless ambiguous utterance tokens, such as strings of filled pauses, that serve only to hold the turn, if anything. For meso-level coding, such "meaningless" tokens are not assigned to CGUs and again are invisible to macro-level analysis. Of course, the role of such tokens in turn-taking or analyzing initiative is meaningful; the treatment of these tokens in the proposed scheme reveals its bias toward directly representing *informational* and *intentional* discourse relations, and not directly representing *initiative* or *turn-taking* phenomena.

Acknowledgments

We would like to thank Masato Ishizaki for helpful comments on previous drafts of this document.

References

[Allen and Core, Draft 1997] James Allen and Mark Core. Draft of damsl: Dialog act markup in several layers. available through the WWW at: http://www.cs.rochester.edu/research/trains/annotation, Draft, 1997.

[Allwood *et al.*, 1989] Jens Allwood, Joakim Nivre, and Elisabrth Ahlsen. Speech management: On the non-written life of speech. Technical Report (GPTL) 58, Gothenburg Papers in Theoretical Linguistics, University of Göteborg, 1989.

[Allwood *et al.*, 1992] Jens Allwood, Joakim Nivre, and Elisabrth Ahlsen. On the semantics and pragmatics of linguistic feedback. *Journal of Semantics*, 9, 1992.

[Bratman, 1987] Michael E. Bratman. *Intention, Plans, and Practical Reason.* Harvard University Press, 1987.

[Clark and Schaefer, 1989] Herbert H. Clark and Edward F. Schaefer. Contributing to discourse. *Cognitive Science*, 13:259–294, 1989. Also appears as Chapter 5 in [Clark, 1992].

[Clark, 1992] Herbert H. Clark. *Arenas of Language Use.* University of Chicago Press, 1992.

[Clark, 1994] Herbert H. Clark. Managing problems in speaking. *Speech Communication*, 15:243 – 250, 1994.

[Dahlbäck and Jönsson, 1998] Nils Dahlbäck and Arne Jönsson. A coding manual for the linköping dialogue model. unpublished manuscript, 1998.

[Dillenbourg *et al.*, 1996] Pierre Dillenbourg, David Traum, and Daniel Schneider. Grounding in multi-modal task-oriented collaboration. In *Proceedings of the European Conference on AI in Education*, 1996.

[Goldman, 1970] A. I. Goldman. *A Theory of Human Action.* Princeton University Press, Princeton, NJ, 1970.

[Gross *et al.*, 1993] Derek Gross, James Allen, and David Traum. The TRAINS 91 dialogues. TRAINS Technical Note 92-1, Department of Computer Science, University of Rochester, July 1993.

[Grosz and Sidner, 1985] B. Grosz and C. Sidner. Discourse structure and the proper treatment of interruptions. In *Proceedings of IJCAI-85*, Los Angeles, Ca., August 1985.

[Grosz and Sidner, 1986] Barbara Grosz and Candace Sidner. Attention, intentions, and the structure of discourse. *Computational Linguistics*, 12(3):175–204, 1986.

[Hobbs, 1996] Jerry R. Hobbs. On the relation between the informational and intentional perspectives on discourse. In Eduard H. Hovy and Donia R. Scott, editors, *Computational and conversational discourse*, pages pages 139–157. Springer Verlag, 1996.

[Kowtko *et al.*, 1991] Jacqueline C. Kowtko, S. Isard, and G. Doherty. Conversational games within dialogue. In *Proceedings of the ESPRIT Workshop on Discourse Coherence*, 1991.

[Lochbaum, 1994] Karen Lochbaum. *Using Collaborative Plans to Model the Intentional Structure of Discourse.* PhD thesis, Harvard University, 1994. Available as Technical Report 25-94.

[Moore and Pollack, 1992] J. Moore and M. Pollack. A problem for rst: the need for multi-level discourse structure. *Computational Linguistics*, 18(4):537–544, 1992.

[Nakatani *et al.*, 1995] Christine H. Nakatani, Barbara Grosz, David Ahn, and Julia Hirschberg. Instructions for annotating discourse. Technical Report 21-95, Center for Research in Computing Technology, Harvard University, Cambridge, MA, September 1995.

[Passonneau and Litman, 1997] Rebecca Passonneau and Diane Litman. Discourse segmentation by human and automated means. *Computational Linguistics*, 23(1):103–140, 1997.

[Pierrehumbert, 1980] J. B. Pierrehumbert. The phonology and phonetics of english intonation. Doctoral dissertation, Massachusetts Institute of Technology, 1980.

[Pollack, 1986] Martha E. Pollack. *Inferring Domain Plans in Question-Answering.* PhD thesis, University of Pennsylvania, 1986.

[Schegloff and Sacks, 1973] Emmanuel A. Schegloff and H. Sacks. Opening up closings. *Semiotica*, 7:289–327, 1973.

[Sikorski and Allen, 1997] Teresa Sikorski and James F. Allen. A scheme for annotating problem solving actions in dialogue. In *Working Notes of the AAAI Fall Symposium on Communicative Action in Humans and Machines*, 1997.

[Traum and Allen, 1992] David R. Traum and James F. Allen. A speech acts approach to grounding in conversation. In *Proceedings 2nd International Conference on Spoken Language Processing (ICSLP-92)*, pages 137–40, October 1992.

[Traum and Heeman, 1997] David R. Traum and Peter Heeman. Utterance units in spoken dialogue. In Elisabeth Maier, Marion Mast, and Susann Luperfoy, editors, *Dialogue Processing in Spoken Language Systems — ECAI-96 Workshop*, Lecture Notes in Artificial Intelligence, pages 125–140. Springer-Verlag, Heidelberg, 1997.

[Traum and Hinkelman, 1992] David R. Traum and Elizabeth A. Hinkelman. Conversation acts in task-oriented spoken dialogue. *Computational Intelligence*, 8(3):575–599, 1992. Special Issue on Non-literal language.

[Traum, 1994] David R. Traum. *A Computational Theory of Grounding in Natural Language Conversation.* PhD thesis, Department of Computer Science, University of Rochester, 1994. Also available as TR 545, Department of Computer Science, University of Rochester.

[Traum, 1998] David R. Traum. Notes on dialogue structure. Unpublished manuscript, 1998.

[Walker, 1993] Marilyn A. Walker. *Informational Redundancy and Resource Bounds in Dialogue.* PhD thesis, University of Pennsylvania, 1993.

A Full Dialogue Examples

A.1 TRAINS Dialogue 91-1.1

```
Total Time: 1'38''      Total Turns: 20      Total Utterances: 63

UU#  Speaker: Utterance
1.1 M: okay
1.2  : I have to
1.3  : ship a boxcar of oranges to Bath by 8 o'clock today
...
2.1 S: okay
3.1 M: um
3.2  : so
3.3  : let's see
3.4  : where are there _oranges_
4.1 S: the oranges are in the warehouse
4.2  : at Corning
5.1 M: oh okay
5.2  : and I see that there's a tanker car there
5.3  : oh we don't want a tanker car do we
5.4  : um
5.5  : I have to get a boxcar
5.6  : to Corning
5.7  : and then I have to load it with oranges and eventually I
  have to get that to Bath
5.8  : by 8 o'clock
6.1 S: right
7.1 M: um
7.2  : ts
7.3  : so
7.4  : what's the cl  /
7.5  : w okay
7.6  : we have to get engines to the boxcars
8.1 S: right
9.1 M:  +um+
9.2  : are there boxcars anywhere except at Bath
```

```
10.1 S: there's boxcars at Bath Danville and Elmira
11.1 M: oh okay
11.2  : ... um
[4sec]
11.3  : and we only need one boxcar
11.4  : okay
11.5  : so
[2sec]
11.6  : aha
11.7  : I see an engine and a boxcar both at Elmira
12.1 S: right
13.1 M: this looks like the best thing to do
13.2  : so we should get
13.3  : ... the eng / engine to picks up the boxcar
13.4  : and head for Corning
13.5  : 's that sound reasonable
14.1 S: sure
14.2  : that sounds good
15.1 M: and from Corning we'll pick up the oranges
15.2  : and um
15.3  : take them to Bath
15.4  : will it / that get m / me
15.5  : do you think that I can get  .. this all over to Bath by 8 o'clock
16.1 S: yeah
16.2  : that gets us to Bath at f / 5 AM
16.3  : so it's plenty of time
17.1 M: okay
17.2  : so
17.3  : um
17.4  : do I tell you what to do at this point
17.5  : [chuckle]
18.1 S: +um+
18.2 S: well
18.3 S: we're done
19.1 M: [chuckle]
20.1 S: +we can now sh ship+
20.2  : we're done with the warmup problem
```

A.2 TRAINS Dialogue 91-8.1

Total Time: 1'26'' Total Turns: 36 Total Utterances: 81

UU# Speaker:Utterance

1.1 M: all right
1.2 : so it's uh
1.3 : it's 12 midnight
1.4 : and we need to get uh
1.5 : a boxcar of oranges to Bath
1.6 : by
1.7 : 8 AM
1.8 : uh
1.9 : so in 8 hours
2.1 S: okay
3.1 M: and uh
3.2 : so are there any oranges
3.3 : at / uh
3.4 : Corning
4.1 S: yes
4.2 : there are
5.1 M: okay
5.2 : is there an _engine_ at Corning
6.1 S: no
7.1 M: all right
7.2 : 's see
7.3 : so
7.4 : why don't we take
7.5 : uh
7.6 : an engine from Elmira
8.1 S: okay
9.1 M: uh
9.2 : are both available
9.3 : there're two engines there I

```
10.1 S: yeah
10.2  : they're both
10.3  : they're both there
11.1 M: okay
11.2  : so
11.3  : wh / why don't we take E2
12.1 S: okay
13.1 M: and uh move it to Corning
13.2  : ... how long will that take
14.1 S: that'll take two hours
15.1 M: okay go ahead and do that
16.1 S: okay
17.1 M: and then um
17.2  : we can take the tanker car
17.3  : in Corning
17.4  : and / uh
17.5  : attach it to the uh
17.6  : the engine E / the engine / engine that's coming over
17.7  : E2
18.1 S: okay
19.1 M: and then send them off to Bath
19.2  : at 2 AM
19.3  : err 2:30 AM
20.1 S: okay
20.2  : it doesn't take any time to
21.1 M: to couple
22.1 S: to couple
23.1 M: okay
23.2  : then send it immediately
24.1 S: so
25.1 M: soon as it gets there
26.1 S:        +although well+
26.2  : sorry
26.3  : it / it takes an hour to load
26.4  : so
27.1 M: okay
28.1 S: we have to do that
```

```
29.1 M: so by th / at 3
29.2  : then how long does it take to get from Corning to Bath
30.1 S: two hours
31.1 M: so fine
31.2  : go ahead
32.1 S:  +so+
32.2  : get to bath
32.3  : at 5 AM
32.4  : so
33.1 M: and then
33.2  : the task is .. finished
34.1 S:             +that should be+  fine
34.2  : yeah
35.1 M: +yeah+
36.1 S: okay
```