Guy Hoffman - Fall 2003

A Reading of
Three Papers on
Collaborative Discourse Theory

D. Litman & J. Allen (1990), "Discourse processing and commonsense plans"

B. Grosz & C. Sidner (1990), "Plans for Discourse"

C. Rich, C. Sidner and N. Lesh (2001), "Collagen: Applying Collaborative Discourse Theory to Human-Computer Interaction"

This report will summarize and critique all three of the above papers, since they are in many ways similar and interrelated, as are my comments on them. Litman and Allen's paper is structurally and logically very close to Grosz and Sidner's article. Rich et al. in turn is in many ways an implementation of ideas originating in Grosz and Sidner (1990).

Litman and Allen are concerned with the relationship between plan recognition in discourse and the underlying commonsense structures that are necessary to support the discourse. Building around a train-trip scenario, they demonstrate the importance of commonsense understanding of plans to the ability to analyze a discourse revolving around this plan. In this capacity they make the distinction between domain plans and discourse plans, where the first are plans specific to the domain discussed and the latter are plans that enable the actual discourse in *any* domain. They proceed to suggest a quasi-practical approach to implement these theoretical constructs in discourse analysis.

In "Plans for Discourse", Grosz and Sidner argue for a similar model based on the linguisitc/intentional/attentional trichotomy put forward in their 1986 paper. They describe mechanisms for plan analysis looking at Discourse Segment Purposes (DSPs), which are the intentions of the current discourse utterance and their relationship to the overall plan and intention. The paper also extends Pollack's SimplePlan model into the collaboration realm ("SharedPlan") by using mutual beliefs and intentions as building blocks.

Finally, in the most application oriented of the three papers, "Collagen" describes a system and API for introducing collaborative discourse theory to user interfaces in order to make them more intelligent. The proposed system uses a focus stack and plan trees based on pre-scripted plan recipes to both analyze the current discourse intentions and generate discourse steps to be useful to the user. Discourse analysis is based on search over these trees within the context set forth by the top of the Focus Stack.

"Collagen" aims to be an application-independent framework, even though it is currently biased towards tutoring applications. The method is roughly based on Grosz and Sidner's theoretical work, albeit much simplified.

As important as collaborative discourse analysis and theory are for human-machine collaboration -- after all, most of our collaborations are heavily based on discussion and our understanding of discussion -- I was not entirely satisfied with the framework laid out in these papers (except for the "Collagen" paper).

It seems that both proposed models go into great depth, but eventually account for little. Too many aspects are left unsaid. With that I mean two things: on one hand many steps to achieve the desired results are trivialized (How do we parse? How do we know what domain to work on? How far up to search for a solution? How do we know that piano-lifting is a collaborative task? How do we distinguish between different user-agent relationships etc.). On the other hand, the models seem so limiting that they end up to not really encompass most or even the core of the desired mechanisms.

In addition was I disappointed to find that even though the domains which these approaches chose to tackle are very limited, the solutions still require quite large leaps of faith (or, in the practical implementation - manual input) to function as autonomous systems. In a way these theories cost much (in terms of functional rigidity) and achieve little (in terms of practical/generalizable applications).

Solving problems in the train domain, for example, a much simpler decision systems might be envisioned as more effective even if aiming for a less lofty goal. On the converse side - these models simplify collaboration to such an extent to render it quite irrelevant to our understanding of human collaboration.

On a more positive note, some of the models proposed, in particular those referring to earlier work, seem useful for building systems. I particularly liked the *prerequisites, effects and decomposition* model quoted in the Litman paper, which seemed to be a useful guideline for plan modeling.

Finally, I did like the "Collagen" paper best of the three. Even though it seemed obvious that it is not as universally generalizing as it claims to be, there are some interesting practical solutions: the Focus Stack is a useful approach and so is "Collagen"'s model of partially ordered hierarchical tasks. The latter is particularly interesting with respect to the collaborative task decomposition work we are planning to do in the near future with Leo.