COLLAGEN: Applying Collaborative Discourse Theory to Human-Computer Interaction

Authors:
Charles Rich, Candace L. Sidner, Neal Lesh

Topic: Collaborative Dialogue

Presenter: Michael A. Sao Pedro

Date: 3/13/2008

Motivation

1 User says "Hello."
2 Agent says "Hello, what's your name?"
3 User says "Candy."
4 Agent says "Hello, Candy!"
5 User says "Let's set the clock."
6 Agent says "What is the time?"
7 User says "The time is 11:00 AM."
8 Agent sets the clock to 11:00 AM.
...
9 User says "How do I play back a video from my camcorder?"
10 Agent says "First, plug in the camcorder cables."
11 User says "How do I plug in the camcorder cables?"
12 Agent points to where you plug in the camcorder cables.
13 Agent says "The red cable plugs into the red terminal, the black cable plugs into the black terminal, and the yellow cable plugs into the yellow terminal. Make sure the other ends of the cables plug into the corresponding colors on the camcorder."
14 User says "What next?"
15 Agent says "Now press Play on your camcorder. You should see a picture on the TV screen."
16 User says "Ok."
17 Agent says "I don't see a picture. Make sure the camcorder is plugged in or has fresh batteries, and is turned on."
18 User says "Ok."
19 Agent says "Good. It's working."

- What makes a UI “intelligent”?
- Mixed initiative (MI) systems do not support user’s problem solving process.
- Common MI limitations:
  - Action order between user and computer inflexible
  - Hard to recover from mistakes
  - Each system has its own interaction conventions
  - Easy to forget already explored options
  - Difficult to recover from getting stuck

Solution: COLLAGEN

- Build better mixed-initiative systems by supporting user’s problem solving process using *collaborative discourse* theory.

- Create a reusable, application-independent collaboration manager for software agents to facilitate full system development.

How do we collaborate?

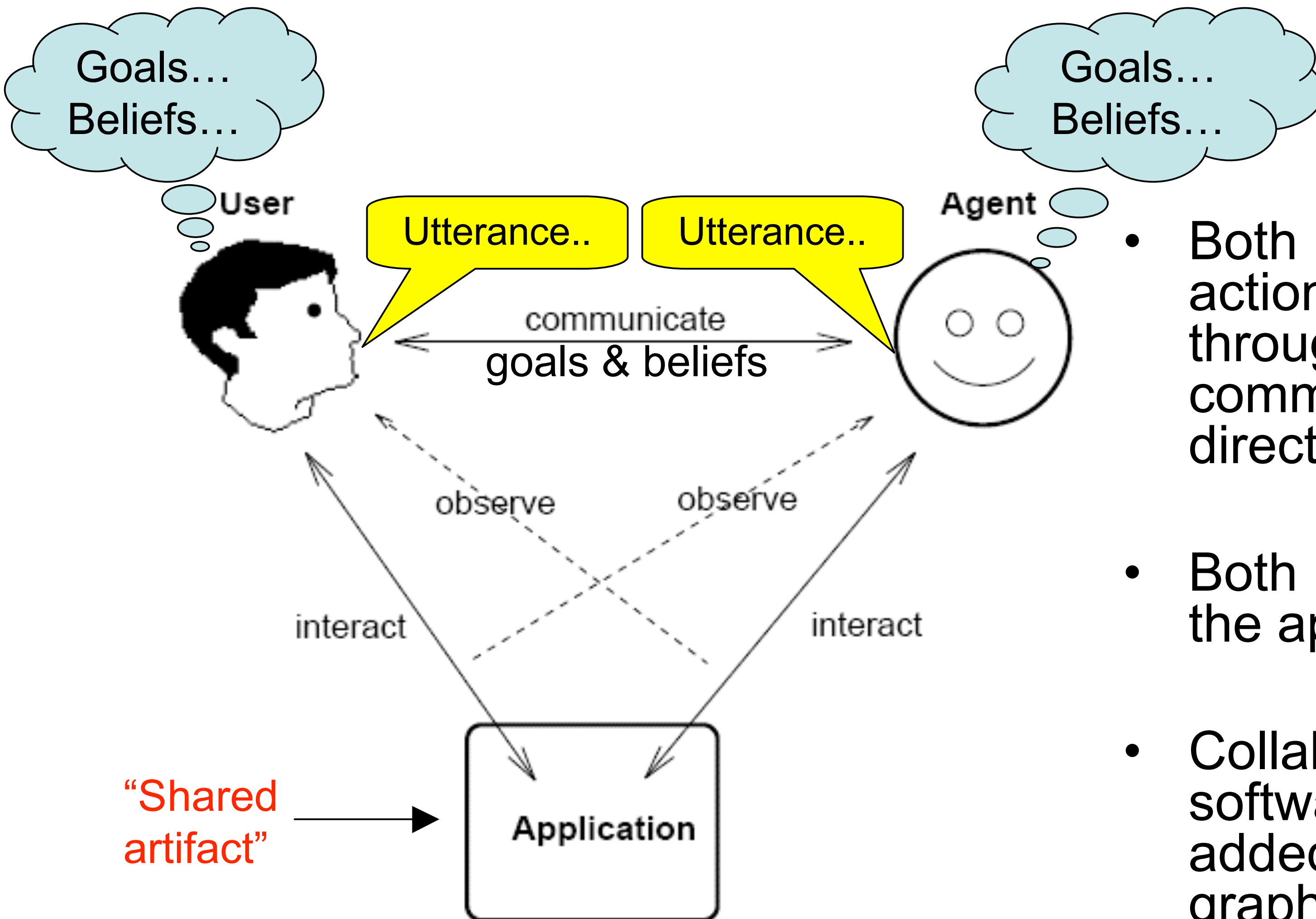
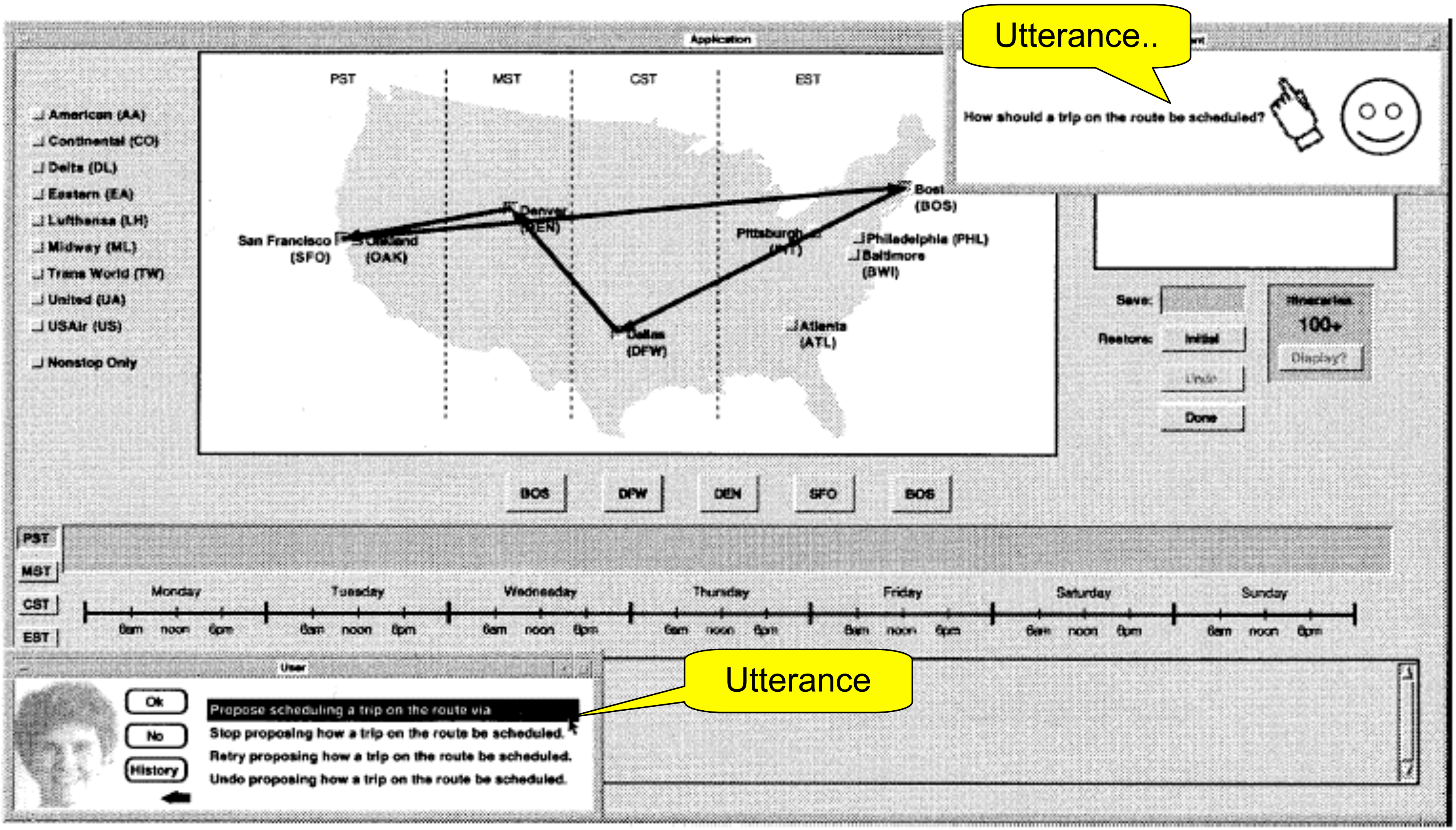


Figure 1. Collaborative interface agent paradigm.

- Both know when action is performed through communication or direct observation.
- Both interact with the application.
- Collaborative software agent is added existing graphical user interface.

Example Interface



Theoretical Model of Task-Oriented Discourse Structure [Grosz and Sidner 1986]

- **Intentional structure**: Accumulating record of mutual beliefs and intentions of participants during a collaboration in a "partial SharedPlan".

- **Attentional structure**: "Focus stack" for handling shifting focus of attention

- **Linguistic structure**: Hierarchical grouping of communication sequences (segments) that serve some purpose.

Intentional Structure: SharedPlan

- Model for collaborative, multi-agent “plans” that captures **beliefs** and **intentions** of **each agent** (human or software) and the whole **group**.
  - Beliefs and intentions are distributed among individual minds.
- “Partial SharedPlan” represents accumulated beliefs and intentions at current point in collaboration.
  - Utterances (communication) and actions (observations) establish “missing” beliefs and intentions.
- Collagen holds beliefs and intentions of <u>exactly *two* agents collectively, not individually</u>.
  - Their “mutual beliefs”

SharedPlan for Two Agents

- Why are SharedPlans complex?
  - Group must have mutual beliefs of a recipe for action.
  - Designate a single agent to perform each subaction in their recipe.
  - Agents must have commitments toward their own and group's actions.

- Two agents "GR" have a SharedPlan for an action "A" if:
  - (1) GR is committed to performing A
  - (2) GR has a *recipe* for A
  - (3) For each singe-agent act B in recipe for A, there is an agent such that
    - (a) B is performed by agent G_b
    - (b) G_b intends to perform B; G_b believes it can perform B; G_b has an *individual plan* for B
    - (c) The group all mutually believes (b) is true
    - (d) The group is committed to Gi's success

Communication: Sidner's Artificial Discourse Language

- Formal, axiomatic representation of semantics of sentences appropriate for collaborative dialogue.
- A "negotiation" of agreeing on goals, actions, and beliefs that comprise planning and acting in the collaboration.
- Series of message types and propositions
  - After message received by agent, certain beliefs and intentions can be regarded as true.
- Example messages (in English) between agents A & B:
  - "A proposes to collaborate on achieving goal G"
  - "B acknowledges A's proposal to achieve G"
  - "B proposes A should perform step Z to achieve G"
  - "A rejects B's proposal to perform step Z to achieve G"
  - "B asks why A performed step Y"
  - "A proposes that the "belt" parameter for step Y should be "this belt"
  - "A tells B that A performed step Y using the 'belt'"

Linguistic Structure (Segmented Interaction History)

TOPLEVEL

Replace Pump and Belt

A: 'Replace the pump and belt please.'

Replace Belt

B: 'OK, I found a belt in the back.'

B: 'Is that where it should be?'
A: 'Yes.'

B: Removes belt.
B: 'It's done.'

Replace Pump

A: 'Now remove the pump.'
. . .
A: 'First you have to remove the flywheel.'
. . .
A: 'Now take off the base plate.'
B: 'Already did.'

Figure 6. Segments in a task-oriented human discourse.

- Segments contain both utterances and actions.
- Each segment is associated with a purpose aka "goal".
- Each purpose contributes to parent segment's purpose.
- Hierarchy mimics the intentional structure hierarchy.

Discourse State Representation

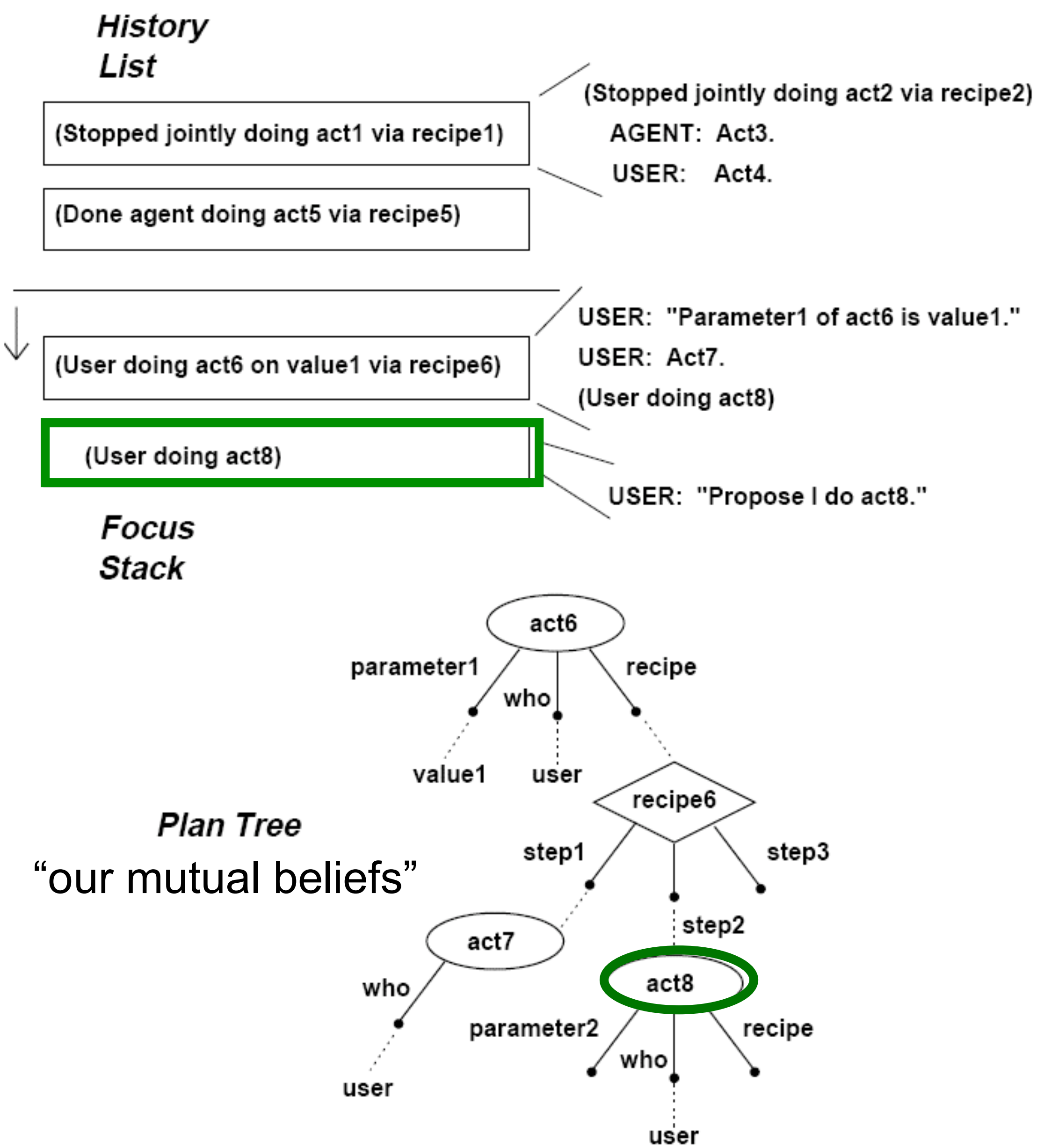


Figure 7. Internal discourse state representation.

- History list contains only toplevel segments.
- Plan trees remain associated with segments even after they are popped off focus stack.

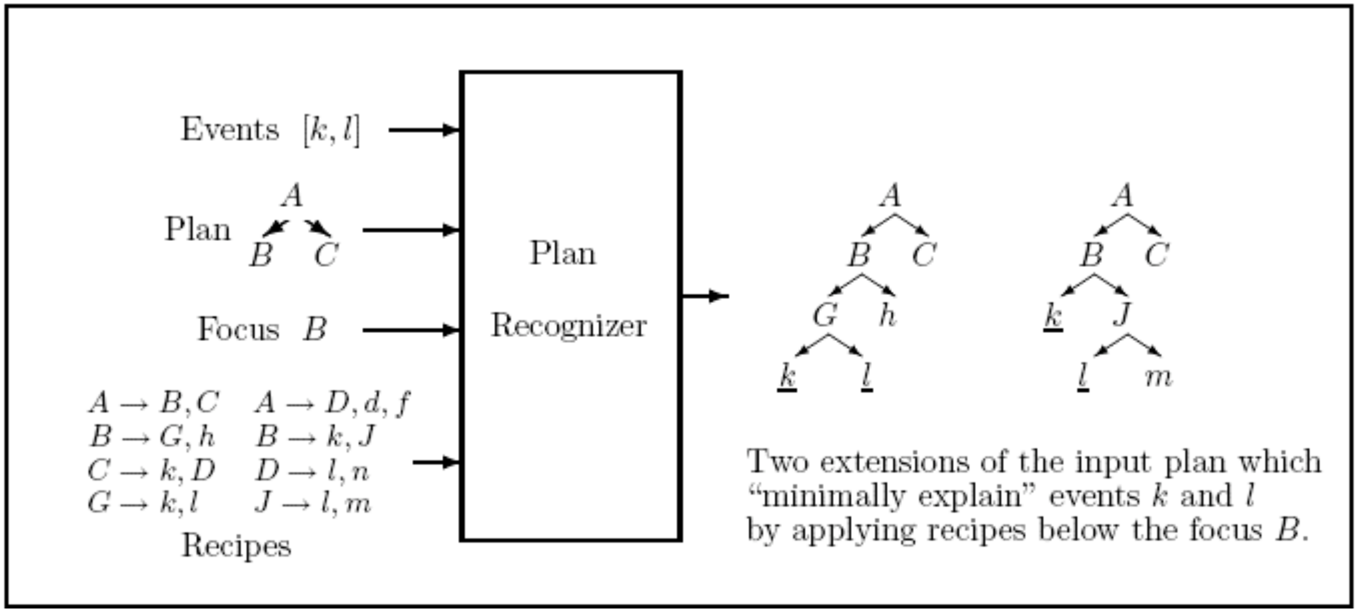
Discourse Interpretation

- Process of updating discourse state after each event generated by human or agent.
  - Utterances + actions = events
  - Does it contribute to the goal in focus or is it an **interruption**?

- Event is relevant to the goal in focus if any are true:
  (1) starts new segment whose goal *contributes* to current purpose
  (2) continues current segment by *contributing* to current purpose
  (3) completes current purpose

- Use plan recognition to see if event can be explained.

Plan Recognition: Inferring Intentions from Actions



- Reduces amount of communication to maintain mutual understanding between user and agent.

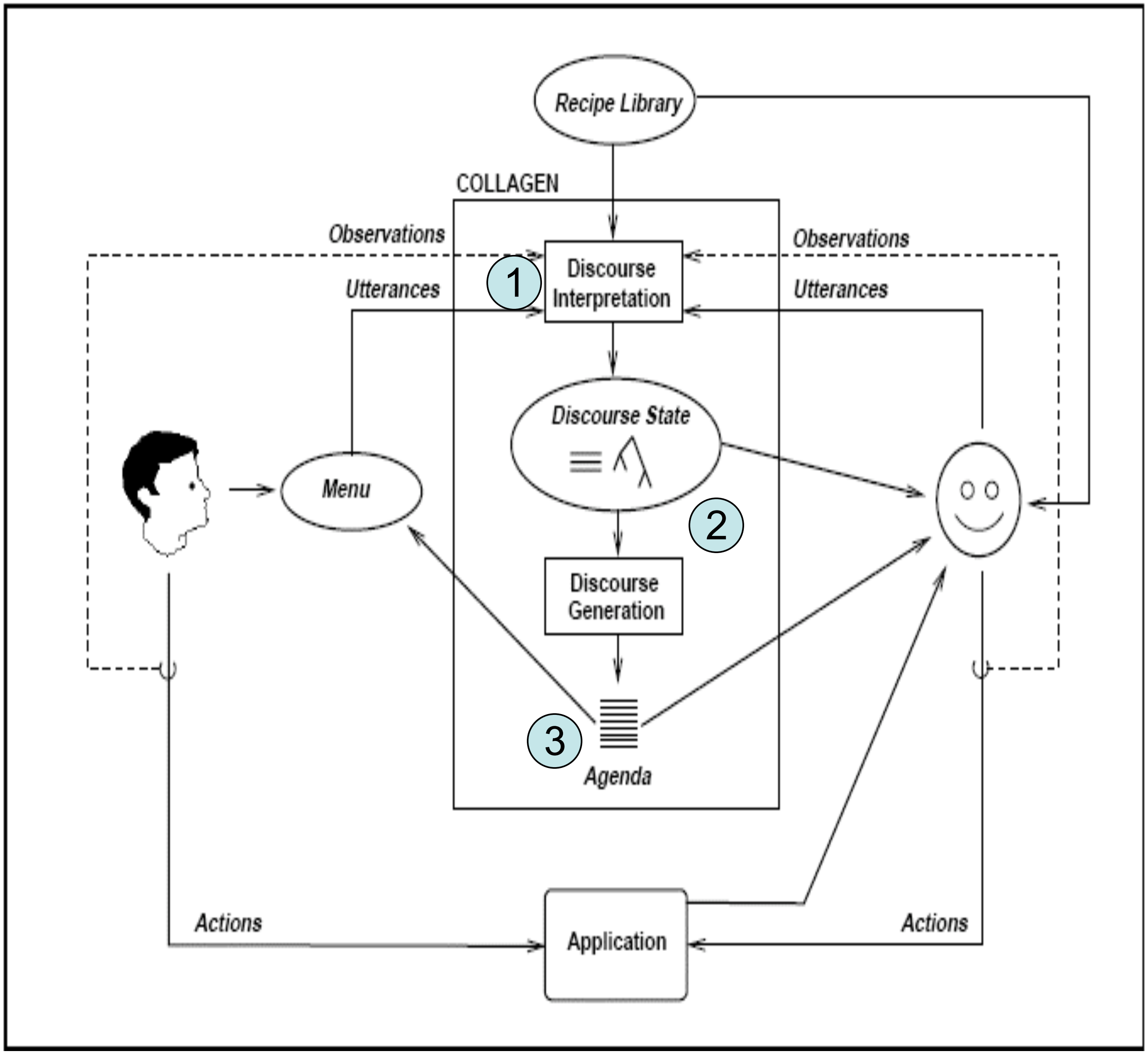
- To make plan recognition tractable:
  - Assume event contributes to purpose at top of focus stack (B)
  - Only incrementally extend a given partial plan
  - If sequence of events gets too large, ask for clarification

Discourse Generation

- Given current discourse state, generate a prioritized list (an agenda) of actions and utterances.

- These can *contribute* to the current discourse goal on the top of the focus stack.
  - Agents can use these to make decisions

COLLAGEN Architecture



- English responses generated by string substitution defined in *recipe library.*

- Developer provides recipe library and API for which application actions can be performed or observed.

Related Work

- Levesque et al. 1990: Different theoretical model of collaboration that does not integrate the intentional, attentional, and linguistic aspects of collaborative discourse.
- Tambe 1997: Collaboration with more than two participants.
- Vermobil project 2000: Reusable across domains; focuses on general linguistic issues in discourse processing, not explicitly modeling collaboration.

Future Work

- Extend plan recognition to relax constraints imposed by recipes
  - Detect performing actions out of order
  - Correct action with wrong parameter
- Explore using Collagen in tutoring domains
  - Detect incorrect action or parameter and let student fix their mistakes.
  - Support “tutorial goals” and recipes encoding “worked examples”
  - PACO: Gas Turbine
- Extensions to Collagen involving discourse theory:
  - Track who is in control of the conversation (attentional)
  - Support negotiations when a proposal is rejected
  - Model agents’ individual beliefs and intentions (great for tutoring)
  - Model more than two agents in collaboration
- Usability testing of Collagenized systems

Contributions and Summary

- Created a reusable, domain independent system that realizes large portions of Grosz and Sidner’s collaborative discourse theory to support building intelligent user interfaces.

- Described how plan recognition could be used in a collaborative setting to support automatically inferring how primitive actions contribute to the plan.

- Gave examples of systems in a variety of domains in which Collagen has successfully been utilized

References

- Rich & Sidner, DiamondHelp: A Generic Collaborative Task Guidance System, AI Magazine 2007

- Rich, C. and Sidner, C. COLLAGEN: A Collaborative Manager for Software Interface Agents. In *User Modeling and User-Adapted Interaction*, Volume 8, Issue 3-4, pages 315-350, 1998.

- Barbara Grosz and Sarit Kraus. 1999. "The Evolution of SharedPlans."   In *Foundations and Theories of Rational Agencies*, A. Rao and M. Wooldridge, eds. pp. 227-262.

- Lochbaum, K. E. A collaborative planning model of intentional structure. *Computational Linguistics* 24(4), 1998.

- Sidner, C. L. An artificial discourse language for collaborative negotiation. In *Proc. 12th National Conf. on Artificial Intelligence*, pages 814-819, 1994.

BONUS MATERIAL!

What Of “Mixed-Initiative”?

- Mixed Initiative systems attempt to support an efficient interweaving of contributions by people and computers for reaching solutions to problems.

- An “integration” of automated services and direct manipulation

- MI systems should:
  – Add value over direct manipulation
  – Maintain working memory of recent interactions
  – Consider status of user’s attention in the timing of a service

ITS’s vs Collaborative Dialog Systems (CDS’s)

- **ITS’s focus on adapting to individual student’s learning needs**
  - **Expert Model**: encoding of how an expert would perform a task
  - **Student Model**: encoding of what a student knows (or doesn’t know)
  - **Pedagogical Model:** encoding of what pedagogical actions to take

- **CDS’s focus on computational models of dialog for collaborative tasks**
  - How to model such interactions in a computer

PACO and Collagen

- **Student (the apprentice) and tutor (the master) work in tandem to learn procedures for operating in a simulation environment.**
  - Student performs actions, asks questions to the tutor
  - Tutor oversees what student does, keeps track of what the student knows, suggests actions, and decides when to let go of the reigns to let the student answer questions.

- **PACO and Collagen combined provide a *domain independent* computational framework that enables the creation of a mixed-initiative learning environment**
  - PACO: Tracks student model and makes pedagogical decisions.
  - Collagen: Maintains discourse state and provides data that PACO uses to make its pedagogical decisions

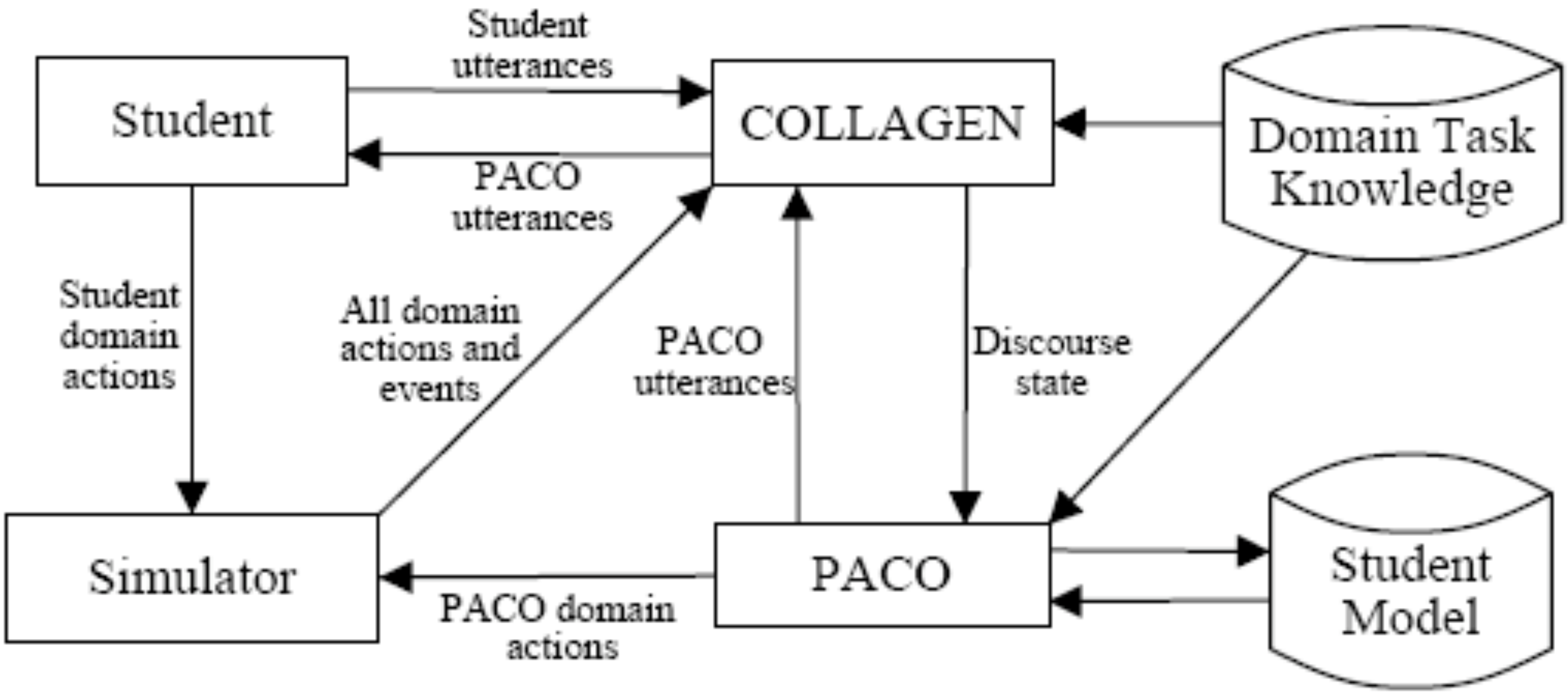


Fig. 2. Paco's Architecture

PACO: Pedagogical Agent for Collagen (USC, Mitsubishi, MITRE)

- **Simulation-based** training
- **Domain-independent**: adapts to any simulator (e.g. Gas Turbine Engine)
- **Collaborative Discourse Theory-based**:
  - Rules describe interactions between three agents: student, tutor, simulator
  - Discourse acts: both utterances and domain actions

Some Math

Collaboration: two or more participants coordinating their actions to achieve a shared goal.

\+ **Discourse**: two or more participants communicating in a shared context

---

Collaborative Discourse: Two or more participants (people and software agents) communicating about how to coordinate achieving a shared goal.

Add'l References

- Eric Horvitz. Principles of Mixed-Initiative User Interfaces. In: Proceedings of CHI '99, ACM SIGCHI Conference on Human Factors in Computing Systems, Pittsburgh, PA, May 1999. ACM Press. pp 159-166