COLLAGEN: Applying Collaborative Discourse Theory to Human-Computer Interaction

Charles Rich
Candace L. Sidner
Neal Lesh

Mitsubishi Electric Research Laboratories
Cambridge, Massachusetts
Road Map

- Philosophy and approach
- Theory and implementation
- Live demos
- Collagen architecture
- Specialized topics
- Related and future work
Collaborative Interface Agent

* SharedPlans per Grosz, Sidner, Lochbaum, Kraus, et al.
COLLAGEN
Java Middleware for COLLaorative Interface AGENts

- air travel planning
- email reading and responding (w. IBM/Lotus)
- GUI design tool operation
- car navigation system operation
- airport landing path planning (w. MITRE)
- gas turbine operator training (w. USC/ISI)
- personal video recorder operation
- programmable thermostat operation (with Delft U.)
- multi-modal web-based form-filling
Let's check our to-do list.
<table>
<thead>
<tr>
<th>Condition</th>
<th>Value</th>
<th>Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>(voltage &gt;= 1000.0000)</td>
<td>default</td>
<td></td>
</tr>
<tr>
<td>(online &amp;&amp; (remote &amp;&amp; controllable))</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(online &amp;&amp; (remote &amp;&amp; !(controllable)))</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(online &amp;&amp; !(remote))</td>
<td></td>
<td></td>
</tr>
<tr>
<td>!(online)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>default</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The diagram shows a symbolic editor interface with options and symbols indicating different conditions and actions.
The Thermostat View interface includes options for living room, kitchen, bathroom, bedroom dave, bedroom lisa, bedroom parents, attic, and study. A user interface with buttons for Yes, Ok, Accomplished, What next?, No, and Stop is also present. The current time is 12:53.
Agent teaching Targets.

Abigail says “Ok.”
Agent says “Hi, Abigail. Welcome to Targets training.”

Agent giving you an exercise to practice building a new RNAV route.
Agent says “Let’s try a route design exercise.”
Agent says “Ok?”
Abigail says “Ok.”

Achieving DomainGoal
Building a new departure route for the Seattle airport by building a new RNAV route using the “new path” tool in TARGETS’ plan view window.
Agent says “Let’s build a new departure route for the Seattle airport.”
Agent says “Every route you design in TARGETS will have its own project containing all the information about the route.”

Done successfully starting defining a new project by Selecting the New Project menu item.
Do successfully entering the properties of the new project.
<table>
<thead>
<tr>
<th>Channel</th>
<th>11 AM</th>
<th>12 PM</th>
<th>1 PM</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABC</td>
<td>The View</td>
<td>NewsCenter 5 at</td>
<td>Port Charles</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Monday</td>
<td>All My Children</td>
</tr>
<tr>
<td>A&amp;E</td>
<td>Magnum, P.I.</td>
<td>Night Court</td>
<td>NewsRadio</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Law and Order</td>
</tr>
<tr>
<td>The Cartoon Network</td>
<td>Dexter's Laboratory</td>
<td>Powerpuff Girls</td>
<td>Scooby Doo Where Are You?</td>
</tr>
<tr>
<td></td>
<td>Scooby Doo Where Are You?</td>
<td></td>
<td>Flintstones</td>
</tr>
<tr>
<td>CBS</td>
<td>The Price Is Right</td>
<td>WBZ 4 News</td>
<td>The Young and the Restless</td>
</tr>
<tr>
<td>CNN</td>
<td></td>
<td>CNN Live Today</td>
<td>Burden of Proof</td>
</tr>
<tr>
<td>Comedy Central</td>
<td>Whose Line Is It Anyway?</td>
<td>Win Ben Stein's Money</td>
<td>Saturday Night Live</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Saturday Night Live</td>
</tr>
<tr>
<td>C-SPAN</td>
<td></td>
<td></td>
<td>U.S. House of Representatives</td>
</tr>
<tr>
<td>The Disney Channel</td>
<td>Bear in the Big Blue House</td>
<td>PB and J Olie</td>
<td>Role Polie Olie</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>The New Adventures of Winnie the Pooh</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Out of the Box</td>
</tr>
<tr>
<td>Entertainment Television</td>
<td></td>
<td></td>
<td>Celebrity Profile</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Coming Attrac</td>
</tr>
<tr>
<td>ESPN</td>
<td>SportsCenter</td>
<td>Baseball Tonight</td>
<td>TBA</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Power of Attorney</td>
</tr>
<tr>
<td>Fox</td>
<td>Divorce Court</td>
<td>Cosby</td>
<td>3rd Rock From the Sun</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Andy Griffith</td>
</tr>
<tr>
<td>USA</td>
<td></td>
<td></td>
<td>sanitized text</td>
</tr>
</tbody>
</table>

**Speech input:**

show schedule

**DiamondTalk Agent:**

Ok, what next?
Let's start the generator. First, press the On button on the generator to start the generator (Please use the mouse to click the button where I am pointing).
Road Map

☑ Philosophy and approach

☐ Theory and implementation

☐ Live demos

☐ Collagen architecture

☐ Specialized topics

☐ Related and future work
Intentional

purposes, contributes

focus spaces, focus stack

Attentional

Linguistic

segments, lexical items

(D Grosz, Sidner, Kraus, Lochbaum 1974-1998)

Java
Implementation

focus stack

purpose tree
Discourse Segments and Purposes

(fixing an air compressor, E = expert, A = apprentice)

E: Replace the pump and belt please.
A: Ok, I found a belt in the back.
A: Is that where it should be?
A: [removes belt]
A: It’s done.
E: Now remove the pump.
...  
E: First you have to remove the flywheel.
...  
E: Now take the pump off the base plate.
A: Already did.

(Grosz, 1974)
Discourse State Representation

Focus Stack

current focus space

replace belt

replace pump and belt

Purpose Tree

replace pump and belt

replace pump

replace belt

E: Replace the pump and belt please.
A: Ok, I found a belt in the back.
A: Is that where it should be?
A: [removes belt]
A: It’s done

(Grosz & Sidner, 1986)
Discourse Interpretation Algorithm

The current (communication or manipulation) act either:

- **starts** a new segment/focus space (**push**)
- **ends** the current segment/focus space (**pop**)
- **continues** (*contributes to*) the current segment/... (**add**)

An act **contributes to the purpose of a segment if it:**

- directly achieves the purpose
- is a *step* in the plan for the purpose *
- identifies the *recipe* used to achieve the purpose
- identifies *who* should perform the purpose or a step in the plan
- identifies a *parameter* of the purpose or a step in the plan

* does **not** include recursive plan recognition (see later topic)

(Lochbaum, 1998)
A Very Simple Example

Recipe Library (Task Model):

```
public top A;

public manipulation B;
public manipulation C;

public recipe R achieves A {
    step B step1;
    step C step2;
}
```
Road Map

- Philosophy and approach
- Theory and implementation
- Live demos
- Collagen architecture
- Specialized topics
- Related and future work
Two Live Demonstrations

(illustrating the wide range of application possibilities)

(1) **Personal Video Recorder**
- home application
- intelligent assistant
- speech only
- push-to-talk using IBM Embedded ViaVoice

(2) **Gas Turbine Operator Training**
- industrial application
- intelligent tutor
- multi-modal (speech and direct manipulation)
- open mic using IBM ViaVoice Pro
Road Map

- Philosophy and approach
- Theory and implementation
- Live demos
  - Collagen architecture
  - Specialized topics
  - Related and future work
Collagen Architecture
Spoken-Language

Recipe Library

Discourse Interpretation

text-to-speech (JSAPI)

speech understanding (JSAPI)

segmented history

Discourse Generation

Agenda

Collagen Architecture
Spoken-Language

Recipe Library

Discourse Interpretation

text-to-speech (JSAPI)

speech understanding (JSAPI)

segmented history

Discourse Generation

Agenda

Collagen Architecture
Spoken-Language

Recipe Library

Discourse Interpretation

text-to-speech (JSAPI)

speech understanding (JSAPI)

segmented history

Discourse Generation

Agenda

Collagen Architecture
Spoken-Language

Recipe Library

Discourse Interpretation

text-to-speech (JSAPI)

speech understanding (JSAPI)

segmented history

Discourse Generation

Agenda

Collagen Architecture
Spoken-Language

Recipe Library

Discourse Interpretation

text-to-speech (JSAPI)

speech understanding (JSAPI)

segmented history

Discourse Generation

Agenda

Collagen Architecture
Spoken-Language

Recipe Library

Discourse Interpretation

text-to-speech (JSAPI)

speech understanding (JSAPI)

segmented history

Discourse Generation

Agenda

Collagen Architecture
Spoken-Language

Recipe Library

Discourse Interpretation

text-to-speech (JSAPI)

speech understanding (JSAPI)

segmented history

Discourse Generation

Agenda

Collagen Architecture
Spoken-Language

Recipe Library

Discourse Interpretation

text-to-speech (JSAPI)

speech understanding (JSAPI)

segmented history

Discourse Generation

Agenda
Specialized Topics

• Artificial discourse language

• Plan recognition reduces explicit communication
Artificial Discourse Language

(1) Formal semantics in terms of beliefs and intentions:

**speaker:** PROPOSE(b)

\[
\text{Believe}(\text{speaker}, b) \\
\text{Intend}(\text{speaker}, \text{Achieve}(\text{speaker}, \text{Believe}(\text{hearer}, b))
\]

**hearer:** ACCEPT(b)

\[
\text{Believe}(\text{speaker}, b) \text{ Believer}(\text{hearer}, b) \\
\text{Believe}(\text{speaker}, \text{Believe}(\text{hearer}, b)) \text{ Believer}(\text{hearer}, \text{Believe}(\text{speaker}, b)) \\
\text{Believer}(\text{speaker}, \text{Believe}(\text{hearer}, \text{Believe}(\text{speaker}, b))) \\
\text{...}
\]

mutual belief

(Sidner, 1994)
(Specialized Topic)
Artificial Discourse Language

(2) Translation to and from natural languages:

speech recognition

“Let’s work on email.”

natural language understanding

PROPOSE(SHOULD(DoEmail(...)))

utterance menu

text to speech

“Let's work on email.”

template substitution

PROPOSE(SHOULD(DoEmail(...)))
Plan Recognition Reduces Explicit Communication

(Lesh, Rich & Sidner, 1999)
without plan recognition

USER:  "Let's work on email."
AGENT:  "You have 5 unread messages."
USER:  "Let's read a message."
USER:  [select message from Chuck]
USER:  "Let's read message in its own window."
USER:  [open message]
USER:  "Let's do something using message."
USER:  [visit http://foo.bar.com]
...long lunch...
USER:  "What was I doing?"
AGEN:  "Working on email.
   Reading message from Chuck.
   Reading message in its own window.
   Done something using message.
   [visit http://foo.bar.com]
   Expecting to react to message (opt)."
USER:  "Let's react to message."
USER:  "Let's send email."
USER:  [start new message]
USER:  "Let's fill in message."
USER:  [address message to Dan]

...additional editing actions...
USER:  "Done filling in message."
AGENT:  "Let's send, cancel or save message."
USER:  "Please save message."
AGENT:  [save message]

with plan recognition

USER:  "Let's work on email."
AGENT:  "You have 5 unread messages."
USER:  [select message from Chuck]
USER:  [open message]
USER:  [visit http://foo.bar.com]
...long lunch...
USER:  "What was I doing?"
AGENT:  "Working on email.
   Reading message from Chuck.
   Reading message in its own window.
   Done something using message.
   [visit http://foo.bar.com]
   Expecting to react to message (opt)."
USER:  [start new message]
USER:  [address message to Dan]
AGENT:  "Why are you sending email to Dan?
   (a) reacting to message from Chuck
   (b) as a new goal"
USER:  "I am reacting to message from Chuck."
...additional editing actions...
USER:  "Done filling in message."
AGENT:  "Let's send, cancel or save message."
USER:  "Please save message."
AGENT:  [save message]
Other Specialized Topics

• Clarification sub-dialogues
• Turn taking and mixed initiative
• Collaborative negotiation
• Near-miss plan recognition
• Learning task models (recipe libraries)
• “Engagement”
Related Work Using Collagen

**Tutoring as a kind of collaboration:**

- Rickel, USC/ISI
- Ohi, Mitsubishi Electric, Japan

**Communicating with intelligent consumer products:**

- Keyson, Delft U. of Technology

**Intelligent assistance for complex software:**

- Gertner, MITRE
- Kojima, Mitsubishi Electric, Japan
Other Related Work

• multiple participant collaboration (vs. two participants)
  e.g., Tambe et al.

• other theoretical models of collaboration (vs. SharedPlan)
  e.g., Levesque & Cohen, Carberry

• application-specific collaborative dialogue systems (vs. middleware)
  e.g., MERIT, MIRACLE, DenK, TRIPS

• other interface agents (without discourse theory)
  e.g., Maes, and many others

• other agent-related middleware (without discourse management)
  e.g., PRS, and other BDI interpreters
Future Work

Other Specialized Topics:

- Clarification sub-dialogues
- Turn taking and mixed initiative
- Collaborative negotiation
- Engagement

www.merl.com/projects/collagen