Agent Conversations and the ADAPT Testbed

NORMAN BADLER
ALEXANDER SHOULSON
CENTER FOR HUMAN MODELING AND SIMULATION
SIG CENTER FOR COMPUTER GRAPHICS
UNIVERSITY OF PENNSYLVANIA
HTTP://WWW.CIS.UPENN.EDU/~BADLER

WITH GREAT APPRECIATION FOR THE WORK OF MANY STUDENTS AND COLLEAGUES...
Goal (Long Term)

Design simulated human-robot team environments that allow interaction between

- Dynamically generated tasks
- Dynamically occurring events
- Dynamically compliant human-bot interactions
Simulating Conversations*

- Define conversational situations
- Define conversational archetypes
- Initiate a conversational event
- Evolve a conversation event

• Classified dyadic conversations into 14 situations dependent on the agents’ spatial relationship.

• Automated spatiotemporal and iterative variations among simulated instances of a conversational archetype.

• Utilized agent attributes, relationships and emotions to influence animations probabilistically.

• Implemented dyadic conversation archetypes as behavior trees.
Dyadic Conversation Model

- Triggering:
  - Environment Context
  - Time
  - Number of Conversations

- Initialization:
  - Conversation Parameter Computation
  - Agent Attributes
  - Environment Context

- Animation:
  - Conversation Archetype
  - Expression
  - Outcome
  - Turn taking
  - Agent Action
  - Agent Gesture
  - Agent Emotion

Behavior Tree:

MURI Annual Review

Oct. 12, 2012
Transitions between Conversation Archetypes

- Simple asking-answering
- Friendly chatting
- Bargaining
- Arguing
Situation Types: Dyadic or Solo

1. Both standing, facing each other
2. Both standing, facing about 45° from forward toward the other agent
3. Both standing, facing the same direction
4. Both walking (or jogging) together (side-by-side)
5. Both walking (or jogging) toward each other and talking very briefly (as in a greeting) “en passant”
6. Both seated, facing each other
7. Both seated, sitting next to each other facing the same (e.g., on a bench)
8. Both seated, facing about 45° from forward toward the other agent
9. One seated and the other standing
10. One seated and the other walking
11. One standing and the other walking
12. One agent is walking using a cell phone
13. One agent is seated using a cell phone
14. One agent is standing using a cell phone
Asking-Answering: Simple Situation Types
## Relate Conversation Archetype and Situation Type

<table>
<thead>
<tr>
<th>Conversation Archetype</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simple asking-answering</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Friendly chatting</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Bargaining</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>Arguing</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
</tbody>
</table>
The distribution of situation types with different environment contexts

<table>
<thead>
<tr>
<th>Environment</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
</tr>
</thead>
<tbody>
<tr>
<td>Street</td>
<td>20</td>
<td>20</td>
<td>5</td>
<td>15</td>
<td>15</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>10</td>
<td>10</td>
<td>--</td>
<td>5</td>
</tr>
<tr>
<td>Office</td>
<td>5</td>
<td>5</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>25</td>
<td>10</td>
<td>10</td>
<td>15</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>Restaurant</td>
<td>5</td>
<td>5</td>
<td>2</td>
<td>5</td>
<td>3</td>
<td>35</td>
<td>20</td>
<td>5</td>
<td>5</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>Crossroad</td>
<td>5</td>
<td>5</td>
<td>10</td>
<td>30</td>
<td>20</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>10</td>
<td>10</td>
<td>--</td>
<td>10</td>
</tr>
<tr>
<td>Marketplace</td>
<td>30</td>
<td>5</td>
<td>5</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>5</td>
<td>5</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>5</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>
The distribution of four conversation archetypes with different environment contexts

<table>
<thead>
<tr>
<th>Environment</th>
<th>Simple asking-answering</th>
<th>Friendly chatting</th>
<th>Bargaining</th>
<th>Arguing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Street</td>
<td>30</td>
<td>60</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Shopping Mall</td>
<td>30</td>
<td>20</td>
<td>45</td>
<td>5</td>
</tr>
<tr>
<td>Marketplace</td>
<td>25</td>
<td>15</td>
<td>50</td>
<td>10</td>
</tr>
<tr>
<td>Restaurant</td>
<td>25</td>
<td>65</td>
<td>--</td>
<td>10</td>
</tr>
<tr>
<td>Crossroad</td>
<td>70</td>
<td>25</td>
<td>--</td>
<td>5</td>
</tr>
</tbody>
</table>

...
Agent Factors

\[ P_{CT2_i} = aD_i + bR + cE + dS + eEP_i \]

\[ P_{CTL_i} = a_1D_i + b_1R + e_1EP_i \]

- \( P \) is probability that 2 (or 1) agents in suitable proximity will have a conversation of a given type (all values in range \([0,1]\)):
  - \( D_i \) represents the distribution of the \( i^{th} \) conversation archetype in this environment context.
  - \( R \) is intimacy between agents: \( 0 = \) strangers; \( 1 = \) close friends.
  - \( E \) is a simple one-dimensional “happiness” (higher is better; it depends on happiness of both agents).
  - \( S \) represents whether one or both agents have another scheduled required event time approaching (=1) or not (=0).
  - \( EP \) represents the estimated probability for the \( i^{th} \) conversation archetype (to insure conversational variety).
- Weights \( a, b, c, d, e \) are set empirically.
The value of $b$ for different inter-agent relationships.

<table>
<thead>
<tr>
<th>Relationship</th>
<th>Simple asking answering</th>
<th>Friendly chatting</th>
<th>Bargaining</th>
<th>Arguing</th>
</tr>
</thead>
<tbody>
<tr>
<td>family members</td>
<td>0.6</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>friends</td>
<td>0.75</td>
<td>1</td>
<td>0</td>
<td>0.1</td>
</tr>
<tr>
<td>coworkers or classmates</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0.1</td>
</tr>
<tr>
<td>buyer-seller</td>
<td>1</td>
<td>0</td>
<td>0.8</td>
<td>0.2</td>
</tr>
<tr>
<td>strangers</td>
<td>1</td>
<td>0.4</td>
<td>0</td>
<td>0.1</td>
</tr>
</tbody>
</table>
Conversation = (ID, CT, ST, R)

- ID ∈ \{1, 2 \ldots\}
- CT ∈ \{1, 2, 3, 4\} = \{simple asking-answering, friendly chatting, bargaining, arguing\}
- ST ∈ \{1, 2, \ldots, 14\} (situations)
- R ∈ \{1, 2, \ldots, 7\} = \{strangers, co-workers, classmates, friends, family members, buyer-seller, waiter-customer\}
ADAPT

- **Agent Development And Prototyping Testbed**
- **C# on Unity 3D game engine**
- Supports numerous character activities
  - Navigation, Path-finding, and Steering
  - Locomotion – Walking and Running
  - Reaching, Gaze Tracking, Sitting/Standing
  - Gesture Animations
- Seamless integration of multiple procedural character controllers
Can leverage both off-the-shelf and in-development character animation technology
- Integration with ICT SmartBody and Recast/Detour library
- Control character activities with Behavior Trees
- Dynamic environments with a virtual populace
- Built for experiments, simulation, and training
- Prelude to higher level narrative, instruction execution and event authoring