Pragbot II Corpus & Mechanical Turk

Oct 2011 Review

Kenton Lee
Purpose

• We were tasked with collecting a corpus of data to study language interaction

• Three years ago, we presented Pragbot 1.0, which collected 100 transcripts
  • Thanks Karl Schultz and Chris Potts

• Last year, we presented Pragbot 2, which was a 3D version created by Chris Czyzewicz
Pragbot 2.1 Review

The game is currently in chat mode. When you and your partner are ready, press the Game Mode button to begin.

**Metrics:**
- Bombs Destroyed: *
- Hostages Rescued: *
- Badguys Killed: *
- Badguys Escaped: *
Pragbot 2.1 gameplay

• Two players coordinating to rescue hostages and destroy badguys
  – Keep Commander alive
  – Keep Junior functioning

• Players have to develop a strategy to keep the badguys from escaping
Pragbot 2.1 environment
Pragbot 2.1 limitations

• Layout of objects in the environment never changed
• Badguy behavior was weakly dependent on the player's actions and not very interesting
• Overall, there was not enough feedback from the environment
• Unstable: Lots of residual bugs
Pragbot 2.2 improvements

• Locations of badguys, hostages and bombs are randomized before every session.
  – Expert effects
  – Data is more generalizable

• Badguys exhibit goal-oriented behavior
  – Actively avoid Commander
  – Aggressively attempt to disable Junior
  – A* path-finding algorithm

• New user interface gives better feedback about how a player's actions alter the environment
  – Score
Pragbot 2.2 user interface

Press ESC to toggle chat mode
Pragbot 2.2 chat interface

<Server> Follow commander's orders and help him destroy the bombs. Good luck.
<Commander> Go defuse the bombs
<Junior> Which way first?
<Commander> left

= ok=

Many sources of feedback contributed to the design (Victoria Schwanda)
Data Collection with Amazon Mechanical Turk

- Mechanical Turk connects requesters (us) and people who wish to be paid to do simple tasks
- Mechanical Turk has been an effective tool for collecting NLP data in the past.
  - Potts, 2009
  - Callison-Burch and Dredze, 2010
- Chris Potts demonstrated that Mechanical Turk can be used for collecting data in game environments.
<table>
<thead>
<tr>
<th>Requester</th>
<th>HIT Creation Date</th>
<th>Expired Date</th>
<th>Time Allotted</th>
<th>Reward</th>
<th>HITs Available</th>
</tr>
</thead>
<tbody>
<tr>
<td>EC Native Speaker</td>
<td></td>
<td>Oct 23, 2011</td>
<td>2 hours 30 minutes</td>
<td>$0.01</td>
<td>4</td>
</tr>
<tr>
<td>Jonathan Hanston</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Requester</th>
<th>HIT Creation Date</th>
<th>Expired Date</th>
<th>Time Allotted</th>
<th>Reward</th>
<th>HITs Available</th>
</tr>
</thead>
<tbody>
<tr>
<td>Line-Jyh Chan</td>
<td></td>
<td>Oct 27, 2011</td>
<td>6 days 23 hours</td>
<td>$0.03</td>
<td>341</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Requester</th>
<th>HIT Creation Date</th>
<th>Expired Date</th>
<th>Time Allotted</th>
<th>Reward</th>
<th>HITs Available</th>
</tr>
</thead>
<tbody>
<tr>
<td>Line-Jyh Chan</td>
<td></td>
<td>Oct 27, 2011</td>
<td>6 days 23 hours</td>
<td>$0.03</td>
<td>318</td>
</tr>
</tbody>
</table>

**Play a video game - Rescue hostages! Transcript will be used for linguistics research.**

<table>
<thead>
<tr>
<th>Requester</th>
<th>HIT Creation Date</th>
<th>Expired Date</th>
<th>Time Allotted</th>
<th>Reward</th>
<th>HITs Available</th>
</tr>
</thead>
<tbody>
<tr>
<td>Michelle Marcus</td>
<td></td>
<td>Oct 27, 2011</td>
<td>6 days 23 hours</td>
<td>$1.00</td>
<td>4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Requester</th>
<th>HIT Creation Date</th>
<th>Expired Date</th>
<th>Time Allotted</th>
<th>Reward</th>
<th>HITs Available</th>
</tr>
</thead>
<tbody>
<tr>
<td>Line-Jyh Chan</td>
<td></td>
<td>Oct 27, 2011</td>
<td>6 days 23 hours</td>
<td>$0.03</td>
<td>343</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Requester</th>
<th>HIT Creation Date</th>
<th>Expired Date</th>
<th>Time Allotted</th>
<th>Reward</th>
<th>HITs Available</th>
</tr>
</thead>
<tbody>
<tr>
<td>Line-Jyh Chan</td>
<td></td>
<td>Oct 27, 2011</td>
<td>24 hours</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

A Survey about Your Basic Mathematical Skills (A 10 seconds questionnaire)

<table>
<thead>
<tr>
<th>Requester</th>
<th>HIT Creation Date</th>
<th>Expired Date</th>
<th>Time Allotted</th>
<th>Reward</th>
<th>HITs Available</th>
</tr>
</thead>
<tbody>
<tr>
<td>Line-Jyh Chan</td>
<td></td>
<td>Oct 27, 2011</td>
<td>24 hours</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

A Survey about Your Basic Mathematical Skills (A 10 seconds questionnaire)
Initial Data Collection

• We collected data from Mechanical Turk over 2 weeks
  – $1 for each game played.
  – Restricted to US participants only, at least 18 years old

• Criteria for payment:
  – Positive score
  – Meaningful communication
  – $0.50 bonus for very rich communication or a good score

• Results:
  – 136 completed sessions
  – 269 non-trivial transcripts
  – Many participants with negative scores had rich interactions
<Commander> Hello

<Junior> hi

<Commander> Are you ready?

<Junior> Yes, I'm ready now.

<Commander> My name is Steve. What's yours?

<Junior> Jim.

<Commander> Cool. Nice to meet you.

<Junior> Nice to meet you too.

<Commander> Let's start by following the right wall.

<Junior> Does that make sense?

<Commander> Ok.

<Junior> Go into the hallway to your right and tell me what you see.

<Commander> What do you see?

<Junior> One bad guy but he disappeared.

<Commander> Any bombs?

<Junior> No, the bad guy flipped me over. I'm stuck.

<Commander> On my way. You followed the right wall correctly?

<Junior> I cleared the bad guy.

<Commander> Well done.

<Junior> I followed the right wall, then I saw the bad guy on the left.

<Commander> Is the room to your right clear?

<Junior> As far as I can see, yes.

<Commander> Continue to the next room on your right.

<Junior> Nothing here.

<Commander> Okay, exit the room and continue right down the hallway.

<Commander> Enter the next room to your right.

<Junior> Report what you see.

<Commander> Hostage and bomb.

<Junior> Is the bomb defused?

<Commander> Yes.

<Junior> Nice work. I will enter and save the hostage.

<Commander> What just happened?

<Junior> I don't know. I was hit.

<Commander> Exit that room and continue to the hallway.

<Commander> When in the hallway, enter the next room on your right.

<Junior> Which way from here?

<Commander> To your right.

<Commander> See anything?

<Junior> Empty.

<Commander> Exit and work the next room to your right.

<Commander> See anything?

<Junior> Hostage is here. I destroyed the bomb.

<Commander> Any other bombs left?

<Junior> No.
Website

• If you would like to try out Pragbot 2.2, you can go to the website and play a game
  – Plus we would love your transcript!
• http://www.seas.upenn.edu/~muri/pragbot2.html
• The initial informed consent form required by IRB.
Future Plans

• Immediate Corpus Collection using MTurk
• Annotation
  • Syntactic
  • Simple Discourse
  • Simple Speech Acts
• New Postdoc at UPenn: Ariel Diertani
  • Syntactic annotator for Penn-Helsinki corpora of Early Modern & Modern British English and the DARPA Ontonotes project.
  • PhD (5/11), UPenn, “Morpheme Boundaries and Structural Change: Affixes Running Amok”
Thanks!

Overviews of MTurk for language research:
