Pragbot II Corpus: Collection and Syntactic Annotation

Oct 2012 Review
Mitch Marcus & Ariel Diertani
Pragbot 2 gameplay

• Two players coordinating to *contain* and destroy badguys
• Rescue hostages, defuse bombs, keep Commander alive, keep Junior functioning
• Players have to develop a strategy to keep the badguys from escaping
Pragbot 2.2 user 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
More on Gameplay

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

• **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 their actions alter the environment**
  – Score
  – Reticle
  – Health indicator
### MTurk: How the participants find us

<table>
<thead>
<tr>
<th>Requester</th>
<th>HIT Expiration Date</th>
<th>Time Allotted</th>
<th>Reward</th>
<th>HITs Available</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jonathan Herston</td>
<td>Oct 25, 2011 (2 days 23 hours)</td>
<td>2 hours 30 minutes</td>
<td>$0.01</td>
<td>4</td>
</tr>
<tr>
<td>Liao-Jyh Chen</td>
<td>Oct 27, 2011 (6 days 23 hours)</td>
<td>24 hours</td>
<td>$0.03</td>
<td>341</td>
</tr>
<tr>
<td>Liao-Jyh Chen</td>
<td>Oct 27, 2011 (6 days 23 hours)</td>
<td>24 hours</td>
<td>$0.03</td>
<td>318</td>
</tr>
<tr>
<td>Mitchell Marcus</td>
<td>Oct 27, 2011 (6 days 23 hours)</td>
<td>24 hours</td>
<td>$1.00</td>
<td></td>
</tr>
<tr>
<td>Liao-Jyh Chen</td>
<td>Oct 27, 2011 (6 days 23 hours)</td>
<td>24 hours</td>
<td>$0.03</td>
<td>343</td>
</tr>
</tbody>
</table>

*Play a video game: Rescue hostages! Transcript will be used for linguistics research.*
Data Collection, Phase I

- Data collected on Mechanical Turk for 2 2-week periods
  - $1 for each game played.
  - Restricted to US participants only, at least 18 years old

- Criterion for payment:
  - Meaningful communication
  - Positive score
  - $0.50 bonus for rich communication or a good score

- Subcorpus:
  - 269 non-trivial transcripts, 136 completed sessions
  - 12,600 words in July, 6,000 words in October
  - Many participants with negative scores had rich interactions
Data Collection, Phase II

• Data collected using a group of gamers
  – Some participants played through M-Turk; others did not
  – Also included regular M-Turkers
  – Session lasted for two hours
  – Still US participants only, at least 18 years old

• More difficult conditions
  – Game was altered so that participants lost if more than one badguy escaped

• Subcorpus:
  – 5000 words produced
  – On average, richer interactions
<Commander> Hello

<Junior> hi

<Junior> ready?

<Commander> Yes im ready now

<Junior> ok

<Commander> My name is steve. whats yours?

<Junior> jim

<Commander> cool. nice to meet you

<Junior> nice to meet you too

<Commander> lets start by following the right wall.

<Commander> does that make sense?

<Junior> ok

<Commander> 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 your

right wall correct?

<Commander> i cleared the bad guy

<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

<Junior> thanks

<Commander> continue to the next room on your right

<Junior> nothing here

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

<Commander> enter the next room to your right

<Commander> report what you see

<Junior> hostage and bomb

<Commander> is the bomb defused?

<Junior> yes

<Commander> nice work i will enter and save hostage

<Junior> what just happened

<Commander> i dont 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 bomb

<Commander> any other bombs left?

<Junior> no
Annotation: Overview

• Approximately 25,000 words
• Conversations are tokenized
  – Orthographical conventions regarding words do not always reflect underlying syntax and must be adjusted
  – Non-standard punctuation was edited
  – Absence of punctuation can lead to ambiguity with regard to sentence boundaries
• Each word is given a tag indicating part-of-speech
• Conversations are parsed
  – Initial parse is done using Bikel parser followed by Gabbard’s null element restoration
  – Parses are then hand-corrected
Annotation: Part-of-Speech Tags

• Transcript:
  – `<Commander>` lets start by following the right wall.

• Part-of-speech tagging
  – Penn Treebank annotation conventions are used
  – All word and punctuation tokens receive a tag
  – Meta-material is tagged CODE

• After pos-tagging:
  `<Commander>`/CODE let/VB s/PRP start/VB by/IN following/VBG the/DT right/JJ wall/NN ./.
Annotation: Parsing

• Parsing output

(S (CODE Commander-184685)
  (NP-SBJ (-NONE- *))
  (VP (VB Let)
    (NP-o (PRP 's))
    (VP (VB start)
      (PP-MNR (IN by)
        (S-NOM (NP-SBJ-o (-NONE- *)))
        (VP (VBG following)
          (NP (DT the)
            (JJ right)
            (NN wall))))))))
  (.. )))
)

• Corrected Parse

(S-IMP (CODE Commander-184685)
  (NP-SBJ (-NONE- *PRO* ))
  (VP (VB Let)
    (S (NP-SBJ (PRP 's))
      (VP (VB start)
        (PP-MNR (IN by)
          (S-NOM (NP-SBJ (-NONE- *PRO*)))
          (VP (VBG following)
            (NP (DT the)
              (JJ right)
              (NN wall))))))))
  (.. )))
)
Future Plans

• Release of first installation of annotated corpus in about a month
  • Sanity checks for errors and inconsistencies underway

• Annotation, Phase II
  • 16,000-17,000 words of raw data ready for annotation
  • Collected in three sessions under the newer conditions

• Collection of new data under new conditions
  • Pending IRB renewal
  • Data collection under simulation of final evaluation conditions & new situations of increased difficulty
Thanks!