Results (update 12/3/2013)
- 9,563 single answers
- 3,035 questions
- 508 players

Future work
- Players can bet (measurement of confidence)
- Different strategies for agreement assessment
- Coarse vs fine-grained sense inventory

Evaluation
- Gold Standard: 115 questions manually annotated by experts
- Several measures of agreement
- Only questions that received exactly 6 answers

Precision and recall based on different agreement measures

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Precision</th>
<th>Recall</th>
<th>F-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relative majority</td>
<td>0.880</td>
<td>0.834</td>
<td>0.857</td>
</tr>
<tr>
<td>Absolute majority (P &lt; 0.5)</td>
<td>0.882</td>
<td>0.782</td>
<td>0.829</td>
</tr>
<tr>
<td>Absolute majority (P = 0.7)</td>
<td>0.945</td>
<td>0.608</td>
<td>0.740</td>
</tr>
<tr>
<td>Unanimity (P = 1)</td>
<td>0.973</td>
<td>0.347</td>
<td>0.512</td>
</tr>
<tr>
<td>Chi-square test (ϕ = 0.05)</td>
<td>0.953</td>
<td>0.521</td>
<td>0.666</td>
</tr>
</tbody>
</table>

When taking into account only confident answers, precision is higher and recall is lower

Conclusion

Future work
- Players can bet (measurement of confidence)
- Different strategies for agreement assessment
- Coarse vs fine-grained sense inventory

Results (update 12/3/2013)
- 9,563 single answers
- 3,035 questions
- 508 players

Game
- Senses is a Game With a Purpose
- Using crowdsourcing for linguistic annotation
- Players are asked to select the senses of nouns and verbs
- Game material is generated from the Groningen Meaning Bank

Purpose
- Goal: to use answers from the game for Word Sense Labeling
  Players assign senses from WordNet 3.1 to words
- Challenge: players are no experts
  We cannot expect consistent, high quality answers
- Solution: check for high agreement on answers
  Consider only answers upon which many players agree

Evaluation

Conclusion

Future work
- Players can bet (measurement of confidence)
- Different strategies for agreement assessment
- Coarse vs fine-grained sense inventory

Results (update 12/3/2013)
- 9,563 single answers
- 3,035 questions
- 508 players

Future work
- Players can bet (measurement of confidence)
- Different strategies for agreement assessment
- Coarse vs fine-grained sense inventory