Acquiring Knowledge about Human Goals from Search Query Logs

Markus Strohmaier, Peter Prettenhofer and Mark Kröll
Motivation

Knowledge about human goals has been found to be important

- goal recognition from user actions (plan recognition)
- the generation of action sequences that implement goals (planning) [Lieberman07]

- How do we define a query containing a user goal?
  - how to lose weight
  - turning blonde hair to dark brown
  - download pictures of angels

Related Research:

- Understanding goals in web search [Broder02, Rose and Levinson 04]
  - high level categorization of queries to improve retrieval on the web
  - yet, we know little about the specific goal instances

- Goals and common sense [Liu and Lieberman04, Lee05]
  - Goal Oriented Search Engine (GOOSE)
  - yet, the acquisition of goals has proven to be difficult
Research Overview

Research Question:
- If and How search query logs can be utilized to overcome the problem of acquiring knowledge about human goals?

Following an exploratory research style, we intend to show:
- contain a small but interesting number of user goals
- Separation by automatic methods
- Difference of search query goals with regard to goals in other corpora (43Things.com)

Expected Results:
- Knowledge about the automatic acquisition of goals out of search query logs
- Knowledge about the nature of goals extracted from search query logs
Results of Human Subject Study(1)

- 4 independent raters
- labeled 3000 queries

<table>
<thead>
<tr>
<th>Pair</th>
<th>$\kappa$</th>
</tr>
</thead>
<tbody>
<tr>
<td>A - B</td>
<td>0.86</td>
</tr>
<tr>
<td>A - C</td>
<td>0.87</td>
</tr>
<tr>
<td>A - D</td>
<td>0.88</td>
</tr>
<tr>
<td>B - C</td>
<td>0.83</td>
</tr>
<tr>
<td>B - D</td>
<td>0.84</td>
</tr>
<tr>
<td>C - D</td>
<td>0.87</td>
</tr>
</tbody>
</table>

Queries Not Containing Goals → Queries Containing Goals
Results of Human Subject Study (2)

Examples:
- bug killing devices
- mothers working from home
- how to lose weight

- Classes appear to be separable
- → Motivates an automatic approach
Experimental Setup

- AOL search query log [Pass06]
- ~ 20 million search queries
- recorded between March 1 and May 31 (2006)
- ethical issues

- pre-processing steps to reduce noise
- → pre-processed set comprises 5,405,547 queries

- labeled queries from the human subject study were utilized as training examples (controversial queries were omitted)
Classification Approach(1)

- **Feature Types**
  - **Set of Words**
    - stop word removal
  - **Part-Of-Speech Trigrams**
    - Maximum Entropy Tagger
    - trained on the Wall Street Journal Corpus

**Example:** Query: “buy a car” → buy/VB a/DT car/NN

Set of words: { buy, car }

Part-of-Speech Trigrams: $ VB DT NN $ → {$ VB DT, VB DT NN, DT NN $}
Classification Approach(2)

- Linear Support Vector Machine [Dumais98]
  - Robust and effective in the area of text classification
  - Weka Machine Learning Toolkit [Witten05]
  - No feature selection

- Performance:
  - 10 trails – 3 fold Cross Validation
  - Values averaged
  - Precision, Recall and F1-Measure for the class: “queries containing goals”

<table>
<thead>
<tr>
<th>Precision</th>
<th>Recall</th>
<th>F1 – Measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.77</td>
<td>0.63</td>
<td>0.69</td>
</tr>
</tbody>
</table>
Result Set

- Applying the learnt classifier results in:
  - Result set containing 118,420 entries
  - 97,454 (82.3%) of them are unique

<table>
<thead>
<tr>
<th>Nr.</th>
<th>Goal Instance</th>
<th>#Users</th>
<th>Nr.</th>
<th>Goal Instance</th>
<th>#Users</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>add screen name</td>
<td>194</td>
<td>11</td>
<td>cancel aol account</td>
<td>46</td>
</tr>
<tr>
<td>2</td>
<td>create screen name</td>
<td>98</td>
<td>12</td>
<td>check my computer</td>
<td>41</td>
</tr>
<tr>
<td>3</td>
<td>rent to own</td>
<td>85</td>
<td>13</td>
<td>skating with celebrities</td>
<td>40</td>
</tr>
<tr>
<td>4</td>
<td>listen to music</td>
<td>78</td>
<td>14</td>
<td>discover credit card</td>
<td>37</td>
</tr>
<tr>
<td>5</td>
<td>pimp my ride</td>
<td>64</td>
<td>15</td>
<td>pimp my myspace</td>
<td>34</td>
</tr>
<tr>
<td>6</td>
<td>pimp my space</td>
<td>61</td>
<td>16</td>
<td>change my password</td>
<td>33</td>
</tr>
<tr>
<td>7</td>
<td>assist to sell</td>
<td>57</td>
<td>17</td>
<td>how to gain weight</td>
<td>32</td>
</tr>
<tr>
<td>8</td>
<td>wedding cake toppers</td>
<td>53</td>
<td>18</td>
<td>enterprise car rental</td>
<td>31</td>
</tr>
<tr>
<td>9</td>
<td>cancel aol service</td>
<td>50</td>
<td>19</td>
<td>manage my account</td>
<td>30</td>
</tr>
<tr>
<td>10</td>
<td>“deleted”</td>
<td>47</td>
<td>20</td>
<td>trick my truck</td>
<td>30</td>
</tr>
</tbody>
</table>
Comparative Evaluation

- Comparison with 43Things.com (Goal Corpus)
  - Social Networking site, where users can share lists of goals they want to achieve
  - Sample of 36,000 entries

- We are interested, if and how the two datasets differ:
  - Nature of goals
  - Scope of goals

- Perform Qualitative Analysis
  - By examining most frequent entries in both data sets (verbs, nouns, goal instances)
Verbs in AOL vs. 43Things

- Top N most frequent verbs of both goal corpora
- Observations:
  - AOL verbs seem to deal with technical issues
  - 43Things contains more verbs reflecting social activity

<table>
<thead>
<tr>
<th>#Verbs</th>
<th>AOL</th>
<th>Overlap</th>
<th>43Things</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>buy, listen, sell, use, play</td>
<td>make, find, get, do, learn</td>
<td>be, go have, read, see</td>
</tr>
<tr>
<td>50</td>
<td>listen, change, look, move, add, remove, clean, install, apply, draw, put, are, set, convert, rent, tell, fix, pimp, wed, check, cook, deleted</td>
<td>get, be, learn, go, make, have, do, read, see, find, buy, take, write, start, stop, eat, want, keep, create, build, play, use, lose, is, grow, know, sell</td>
<td>s, become, meet, finish, live, watch, run, give, spend, try, own, improve, love, organize, save, speak, join, visit, attend, ride, let, work, am</td>
</tr>
</tbody>
</table>
Observations:

- AOL goals seem to deal with health related issues:
  
  be anorexic, be bulimic, be emo

- 43Things goals appear to exhibit a more positive sentiment:
  
  be loved, be debt free, be healthy
Observations:

- Users seem to have different time frames in mind. AOL Users often appear to seek immediate answers:
  
  *get rid of ants, get out of debt, get rid of moles*

- 43Things Users do not seem to underlie these time constraints:
  
  *get in shape, get up earlier, get bachelor degree*
Contribution

- automatically extracting user goals from search queries seems feasible to a certain extent

- examination of query instances rather than a high level categorization [Broder02]

- search query logs appear to be a promising resource to acquire human goals in an automatic way as opposed [Lieberman07] where human engagement is required
Discussion

- Does the knowledge about a user's search intent allow an improvement concerning the retrieval task?
  - only a very small percentage of queries contain user goals
  - [He07] already attempt to predict user goals from search queries

- Is it likely that users are going to change their attitude about expressing their latent search intent in an explicit way.
Thanks for your attention!
Questions and Discussion
References(1)


References (2)
