Automatic Extraction of Human Activity Knowledge from Method-Describing Web Articles

May 19, 2010

Jihee Ryu, Yuchul Jung, Kyung-min Kim, Sung-Hyon Myaeng

IR & NLP Lab
Computer Science Department
KAIST
Motivation

- People need helps around their activities (context)
- Intelligent systems need to provide customized services

User information:
- On U.S. highway
- 1 year driving experience
- Heading to New York
- Graduate student

- Loosen lug nuts on tire
- Install spare tire
- Call AAA
- Explain the situation
Motivation

- User feedback
- Realizing that the bank balance is too low to call the emergency road service, she decides to do it by herself and clicks on “Change a Flat Tire When You Are a Woman Alone” (a kind of feedback)
Motivation

- To understand what the user is up to, we need to
  - infer goal activity from the contextual info
  - know actions to be taken for the goal

- Activity knowledge is not available in large quantity
  - Most automatically acquired KB are entity-oriented
    - e.g. concept hierarchy or ontology
  - Activity knowledge is dynamic in nature
    - Difficult to acquire knowledge about daily lives of people

- Past attempts
  - Linguistic knowledge (e.g. WordNet)
  - Narrowly defined area of human activities
  - Comprehensive KB is labor intensive (e.g. CYC)
  - Often mixed with other types of general knowledge (e.g. ConceptNet)
Common Sense Knowledge

- **Open Mind Common Sense (OMCS) (2002)**
  - More than 729,000 raw sentences representing commonsense knowledge
  - Collected from the general public through a template-based Web interface

- **EventNet (2005)**
  - Based on OMCS, about 30,000 temporal relation instances extracted from commonsense events
  - Utilized as a basis for an intelligent household system
    - Knowledge is skewed toward household activities
Activity Extraction

- Shah and Gupta (2005)
  - Based on Open Mind Indoor Common Sense (OMICS)
    - Composed of task steps to populate an initial robot knowledge base
    - For default household task plans
  - A simple activity extraction method
    - Part-of-speech (POS) tagger is applied
    - The first verb and noun phrase extracted as the action and its object
    - Used as a baseline for our experiment
Activity Modeling

- Perkowitz et al. (2004)
  - To sense and model activities from the real world
  - Scale to a much larger class of activities than before
  - Activity extraction
    - Sequence of objects that are physically involved in an activity and their probabilities
    - Probabilistic translations from terms that represent an activity name to terms that represent the objects involved
      - E.g. drinking tea ➔ teacup, teabag, …
**Use of Web Resources**

**e-how** ([http://www.ehow.com/](http://www.ehow.com/))
- A collection of method-describing articles

<table>
<thead>
<tr>
<th>Title</th>
<th>How to Make Omelet Soup</th>
</tr>
</thead>
</table>
| Step 1 | Place the water or canned chicken broth in a large saucepan.  
Boil the sweet yellow onion for several minutes. |
| Step 2 | Add the powdered chicken broth along with the canned mushrooms.  
Boil the soup for a few more minutes, and then add the chopped green onion. |
| Step 3 | Drop the eggs into the simmering broth a few minutes before you're ready to serve the omelet soup. |
# e-how Statistics

<table>
<thead>
<tr>
<th>Category</th>
<th># Articles</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arts &amp; Entertainment</td>
<td>68,165</td>
<td>6.7%</td>
</tr>
<tr>
<td>Business</td>
<td>31,846</td>
<td>3.1%</td>
</tr>
<tr>
<td>Careers &amp; Work</td>
<td>39,291</td>
<td>3.9%</td>
</tr>
<tr>
<td>Cars</td>
<td>30,900</td>
<td>3.1%</td>
</tr>
<tr>
<td>Computers</td>
<td>47,450</td>
<td>4.7%</td>
</tr>
<tr>
<td>Culture &amp; Society</td>
<td>26,508</td>
<td>2.6%</td>
</tr>
<tr>
<td>Education</td>
<td>30,677</td>
<td>3.0%</td>
</tr>
<tr>
<td>Electronics</td>
<td>18,876</td>
<td>1.9%</td>
</tr>
<tr>
<td>Fashion, Style &amp; Personal Care</td>
<td>49,270</td>
<td>4.9%</td>
</tr>
<tr>
<td>Food &amp; Drink</td>
<td>75,842</td>
<td>7.5%</td>
</tr>
<tr>
<td>Health</td>
<td>122,152</td>
<td>12.1%</td>
</tr>
<tr>
<td>Hobbies, Games &amp; Toys</td>
<td>74,216</td>
<td>7.3%</td>
</tr>
<tr>
<td>Holidays &amp; Celebrations</td>
<td>22,632</td>
<td>2.2%</td>
</tr>
<tr>
<td>Home &amp; Garden</td>
<td>102,843</td>
<td>10.2%</td>
</tr>
<tr>
<td>Internet</td>
<td>24,938</td>
<td>2.5%</td>
</tr>
<tr>
<td>Legal</td>
<td>9,805</td>
<td>1.0%</td>
</tr>
<tr>
<td>Parenting</td>
<td>19,427</td>
<td>1.9%</td>
</tr>
<tr>
<td>Parties &amp; Entertaining</td>
<td>8,874</td>
<td>0.9%</td>
</tr>
<tr>
<td>Personal Finance</td>
<td>41,086</td>
<td>4.1%</td>
</tr>
<tr>
<td>Pets</td>
<td>30,017</td>
<td>3.0%</td>
</tr>
<tr>
<td>Relationships &amp; Family</td>
<td>25,220</td>
<td>2.5%</td>
</tr>
<tr>
<td>Sports &amp; Fitness</td>
<td>74,930</td>
<td>7.4%</td>
</tr>
<tr>
<td>Travel</td>
<td>29,359</td>
<td>2.9%</td>
</tr>
<tr>
<td>Weddings</td>
<td>8,449</td>
<td>0.8%</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>1,012,773</strong></td>
<td><strong>100.0%</strong></td>
</tr>
</tbody>
</table>

*Update: Aug. 26, 2009 (1 million articles)
*Review: May 19, 2010 (1.5 million articles)*
Our Work

- **Observation**
  - 24 categories cover a variety of task domains.
  - A sequence of actions associated with a goal (how to …) can be considered a solution to the task.

- **KB Construction**
  - Title ➔ Goal
  - Steps in a how-to instruction ➔ Actions

  **Activity model = (G, A, I)**
  
  - G : Goal
  - A : Action sequences
  - I : Ingredients (objects, location, time)

- Form a hierarchical activity network
How to Make Omelet Soup

**Step 1)** Place the water or canned chicken broth in a large saucepan. Boil the sweet yellow onion for several minutes.

**Step 2)** Add the powdered chicken broth along with the canned mushrooms. Boil the soup for a few more minutes, and then add the chopped green onion.

**Step 3)** Drop the eggs into the simmering broth a few minutes before you're ready to serve the omelet soup.

*how-to article*
Activity Knowledge Base (recap)

- Three Major Elements
  - Activity (or goal)
    - Extracted from the title of a how-to article
  - Action Sequences
    - Specific steps to be taken for the given task in the article
  - Ingredients (location, time, object)
    - Entities associated with action steps

- Network of goals, actions, and ingredients [JWS 2010]
  - Actions can be linked to multiple goals
  - Normalization
Preprocessing

- Sentence boundary detection
  - Identification of imperative sentences

- Parsing
  - Using Stanford NLP library
    - E.g. *loosen the nut* \(\rightarrow\) (VP (VB loosen) (NP (DT the) (NN nut))

- Dependency tree generation
  - Simplify parse trees
    - Eliminate adverbial phrase, determiners, and articles
    - E.g. *also contact them using a phone* \(\rightarrow\) *contact them using phone*
  - Convert a parse tree to a dependency structure
    - E.g. (VP (VB start) (PRT (RP up)) (NP (NN car)))
      \(\rightarrow\) prt(start-1, up-2)
      obj(start-1, car-3)
Target Extraction

- To extract (Goal, Actions, Ingredients) from each article

- Syntactic Pattern-based Method
  - Obvious patterns ➔ high precision
  - But coverage is limited.

- CRF-based Method
  - For more complete coverage
Syntactic Pattern-based Method

- **Pattern discovery**
  - Mask words ➔ generate dependency relation driven patterns
    - E.g. prt(start-1, up-2) & dobj(start-1, car-3) ➔ prt(a, b) & dobj(a, c)
  - Identify frequent patterns (f ≥ 3)
    - E.g. prt(a, b) & dobj(a, c) ➔ verb(‘a b’) / ingredient(c, ‘a b’)
  - Compute confidence for each pattern
    - verb(‘a b’) / ingredient(c, ‘a b’) [confidence: 95%]
    - 184 patterns were generated (confidence > 85%)

- **Instance generation**
  - Extract (action, ingredient) instances based on pattern matching
    - E.g. check out the engine ➔ (Action: check out, Ingredient: engine)
CRF-based Method

- To extract targets not matching known patterns
  - Increase coverage
- General and expressive modeling technique to label objects

Training data
- Take up to 5 sentence instances extracted by applying each of the selected patterns
Feature Generation

- Feature candidates
  - POS with order
    - E.g. you tighten nut → you(PR1) / tighten(VB1) / nut(NN1)
  - Dependency type for word pairs
    - E.g. you tighten nut
      → you(nsubjB1) / tighten(nsubjA1, dobjA1) / nut(dobjB1)
  - Phrase chunk with order (Open NLP tool)
    - E.g. you tighten nut
      → you(NP1) / tighten(VP1) / nut(NP2)

- POS and dependency features were used finally.
Evaluation Overview

- **Purpose**
  - Measure performance of proposed automatic extraction methods

- **Method**
  - A manually constructed test collection
    - 2400 eHow articles from 24 domains were randomly chosen.
    - 10 graduate and undergraduate students annotated the data for actions and ingredients
      - A tool was provided.
    - 2 judgments for each sentence
  - Comparison with baselines (previous approaches)
Annotation Tool

Document Annotation System

Guideline | Example | Workbench | Resource | Review | Feedback
---|---|---|---|---|---

# Document Annotations

Document: Culture & Society/0511_How to Apply for a Commercial Loan for a Church.txt

Currently activated document

Title: How to Apply for a Commercial Loan for a Church

Subtitle: How to Apply for a Commercial Loan for a Church

Step 1: Ensure that the church already has a governing board in place.

Verb form: Ensure

Ingredient item (noun clause): that the church already has a governing board in place

Range button
Save button

Compile a list of the names for all individuals on the board of directors.

Verb form: Compile

Ingredient item (noun phrase): a list of the names for all individuals on the board of directors

Range button
Save button

The list should also contain the profession of each board member, along with their addresses.

Range button
Save button
Result

- Comparisons with baselines
  - Baseline 1
    - Extract every first verb and first noun phrase as an action
  - Baseline 2
    - Extract every first verb and every noun phrase under 'object' and 'substance' categories in WordNet

<table>
<thead>
<tr>
<th>Method</th>
<th>Average Accuracy</th>
<th>Average Coverage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline 1 (based on Shah and Gupta)</td>
<td>0.7866</td>
<td>0.9821</td>
</tr>
<tr>
<td>Baseline 2 (based on M. Perkowiz et al)</td>
<td>0.5432</td>
<td>0.9897</td>
</tr>
<tr>
<td>Syntactic Pattern-based Method</td>
<td>0.9130</td>
<td>0.5660</td>
</tr>
<tr>
<td>CRF-based Method</td>
<td>0.8192</td>
<td>0.9499</td>
</tr>
<tr>
<td>Pattern-based &amp; CRF-based</td>
<td>0.8261</td>
<td>0.9501</td>
</tr>
</tbody>
</table>
Limitations

- Prepositional phrases are not processed
  - Loss of essential information about ingredients
  - E.g. *check for dates on the inside of the wheels* ➔ *(check for, dates)*

- Multi-word expressions are not handled properly.
  - E.g. *pay attention to words of hymn* ➔ *(pay, attention)*

- Limited verb scope handling
Conclusion

- Proposed a method for constructing a large-scale human activity KB (~.5 activities & 3.2M actions) by using how-to articles on the Web
- Need further applications of NLP techniques for context-providing sentences, verb scoping, & prepositional phrases
- The constructed KB is to be complemented with mined experiences from blogs [ACL 2010]
Action Normalization

- Construct an equivalence class of actions with a representative name ➔ clustering

Goals

<table>
<thead>
<tr>
<th>Change a flat tire</th>
<th>Fix a flat tire</th>
</tr>
</thead>
<tbody>
<tr>
<td>(pump, brake pedal)</td>
<td>(pump, brake foot pedal)</td>
</tr>
<tr>
<td>(take, spare tire)</td>
<td>(check, equipments)</td>
</tr>
<tr>
<td>(jack up, car)</td>
<td>(raise, vehicle)</td>
</tr>
</tbody>
</table>

additional Descriptor

contextual Similarity

synonyms

Mapped into same cluster
Reference


