Leveraging Procedural Knowledge for Task Understanding

Zi Yang, Eric Nyberg
Language Technologies Institute
Carnegie Mellon University
Overview

• Reuse an earlier proposed method.
  – It only used wikiHow corpus (freeBaseID ignored).

• Hypotheses
  – wikiHow has enough coverage of the topics?
  – Subtasks suggested by the articles are enough for the task accomplishment?
  – The proposed keyphrase extraction method satisfies the task objective?
Approach

• Identify a task from wikiHow
  – Use Lucene’s default setting to index and retrieve top 5 articles → Lucene score
  – Keep or ignore “stub” articles

• Generate candidate key phrases from “steps”
  – A pretrained CRF-based model (annotate wikiHow’s titles in the full text content) → Likelihood score
  – Keyphrases tend to be VP

• Rank the candidates
  – Different score combination methods
Result & Analysis: Task Retrieval

getting organized at work

– Get Organized and Concentrate on Your Work (4.082339)
– Be Organized at Work (3.4436269)
– Organize School Work (2.709037)
– Organize Work Space (2.6684413)
– Organize Work for School (2.5773914)
Result & Analysis: Task Retrieval

elliptical trainer

– Work out on an Elliptical Machine (2.6765485)
– Use an Elliptical Machine (2.5518372)
– Use Elliptical Exercise Equipment Effectively (2.4160774)
– Draw an Elliptical Arch (2.2027557)
– Make Elliptical Garden Beds (2.0584767)
Result & Analysis: Task Retrieval

choose bathroom

- Choose Lights for a Bathroom (4.8557177)
- Choose the Right Bathroom Vanities (4.751217)
- Choose Bathroom Towel Colors (4.6867957)
- Choose the Right Shower Curtain for Your Bathroom (3.5411198)
- Redecorate a Bathroom (2.2694883)

All required subtasks for “choosing bathroom”
Result & Analysis: Task Retrieval

task ambiguity for NP topics

elliptical trainer

– How to buy
– How to use
Result & Analysis: Task Retrieval

**statistics**

- The same task is identified (19)
- Subtask is identified (19)
  - May be due to ambiguity
- Super-task is identified (10)
  - Topics contain constraints: raise venture capital ca
- No relevant task is identified (2)
Result & Analysis: Keyphrase Extraction

a good example

how to cook pork tenderloin

– buy fresh meat
– kick grilled tenderloin
– cook it
– chill meat
– slice your tenderloin
Result & Analysis: Keyphrase Extraction

more bad examples

- buy one
- comfortable
- buy
- sell
- expert ”’ elevation
- look it up online

short, semantically incomplete
Improvement Ideas

• Test other keyphrase extractors (supervised, unsupervised, combination, etc.)
• Candidate keyphrase ranking
• Additional task/procedure/etc. knowledge base
THANK YOU