

InteractiveMediaMINE at the NTCIR-11 IMine Search Task

Graduate School of Engineering, Kogakuin University
Shohei MINE

Faculty of Informatics, Kogakuin University
Takuma MATSUMOTO, Tomofumi YOSHIDA,
Takuya SHINOHARA, Daisuke KITAYAMA

Background

- TaskMine aims to explore the methods of automatically finding subtasks of a given task
- For instance
For a given task “lose weight,” the possible outputs can be “do physical exercise,” “take calories intake,” “take diet pills” and so on

Outline

- Our method consists of four steps
 - **Extending Query**
 - **Retrieving**
 - **Extracting Tasks**
 - **Ranking**
- We use Yahoo! Chiebukuro(web-based Q&A service) as our system resource
 - In general, Q&A services aim to collect answers that solve the user's problems
 - We expect Yahoo! Chiebukuro to be useful as our system resource for mining tasks

Framework of Our Method

1. Extending Query

- We extend the query entered by the user in order to optimize the search engine

2. Retrieving

- We use Yahoo! Chiebukuro for the extended query and retrieving the top 10 pages of the search results

3. Extracting Tasks

- We extract candidates of tasks from “Yahoo! Chiebukuro”

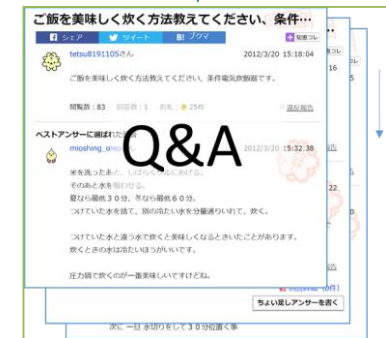
4. Ranking

- We calculate the score of the extracted tasks by the words frequency of each sentence

(Extracted tasks are in the descending order of the score)

Query: “ご飯を炊く”

ご飯 炊く 方法



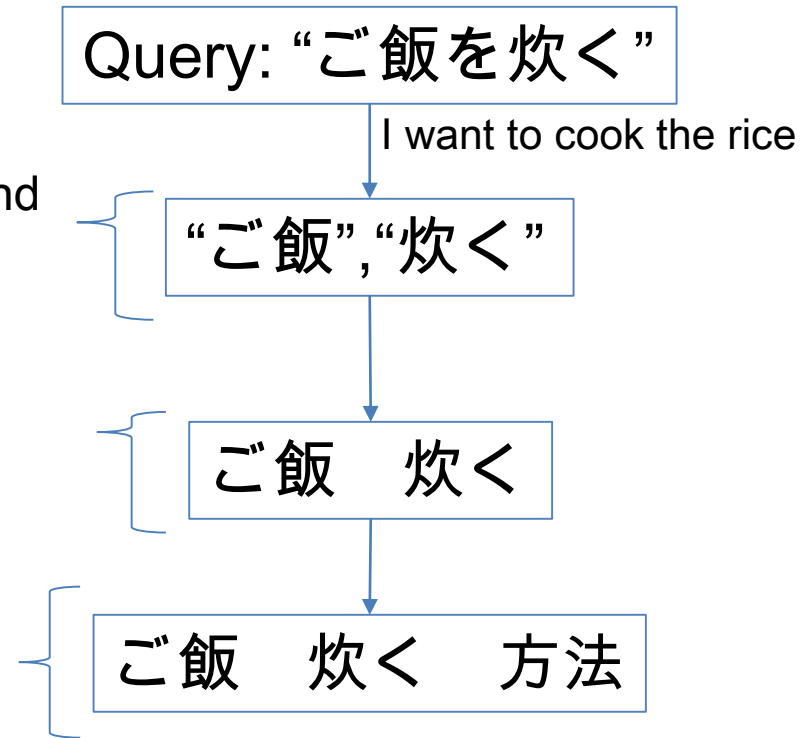
お米2合を普通にといで鍋に入れて水を2カッププラス大さじ2杯入れる

切ったお水も切ります。鍋が温かくなると、お水も切ります。つけていたお水も、鍋の底に少し残しておいて、炊く。つけていたお水と一緒で炊くと美味しくなると思います。炊くとお水は抜きたいほうがいいです。圧力鍋で炊くのが一番美味しいですね。

| Rank | Candidate tasks | NounFreq | VerbFreq | Score |
|------|--|----------|----------|-------|
| 1 | 鍋を中火にかけて沸騰すれば1~2分間キープして直ぐに火を最低限まで絞って15分間、最後に一瞬強火にして火を止める | 17 | 1 | 18 |
| 2 | 炊き方は、洗った米を“ザル”にあげて水を切って“30分したら、分量の水を加える | 14 | 2 | 16 |
| 3 | 言われるように、鍋に米を入れて指の第一関節や手首までお水を入れる | 14 | 2 | 16 |
| 4 | お米2合を普通にといで鍋に入れて水を2カッププラス大さじ2杯入れる | 12 | 2 | 14 |
| 5 | 米を炊く時は、基本的に吸水させたあとに、最大火力になるんですが、火を止める | 8 | 2 | 10 |

Extending Query

1. Our system executes a morphological analysis for the inputted query by user, and extracts nouns and verbs from the query
2. Our system joins the extracted nouns, a single-byte blank, and verbs to make a sentence
3. We add “方法(hoho)” at the end of the query



※方法 is a word that means “method” or “way” in Japanese,
And it is effective in retrieving pages that include questions about
methods used for solving certain problems

→Next we search this query

Retrieving

ご飯 炊く 方法

1. Our system uses Yahoo!
Chiebukuro with the extended
search query

ご飯をおいしく炊く方法 ご飯をおいしく炊く方法ってどんなものがあります...
方法ってどんなものがありますか? 「水を量える系」「米の中何か入れる系」「あきらめる系」なんでも知りたいです。
解決済み - 更新日時:2010/10/04 05:52:39 - 回答数: 5 - 閲覧数: 285
暮らしと生活ガイド > 料理_グルメ、レシピ > レシピ

ご飯を最高に美味しく炊く方法を教えてください。例えば、電気炊飯器、ガス...
ガスがま、土鍋など、ご飯の美味しさを最高に引き出す炊き方を教えてください。
解決済み - 更新日時:2013/10/12 06:48:59 - 回答数: 2 - 閲覧数: 114
暮らしと生活ガイド > 料理_グルメ、レシピ > 料理_食材

ご飯を美味しく炊く方法教えてください、条件電気炊飯器です。
米を洗ったあと、しばらくザルにあげる。そのあと水を吸わせる。夏なら最低30分、冬なら最低60分。つけていた水を捨て、別の冷たい水を分量通りいれて、炊く。つけていた水と違う水で炊くと美...
解決済み - 更新日時:2012/03/22 17:58:05 - 回答数: 1 - 閲覧数: 82
暮らしと生活ガイド > 料理_グルメ、レシピ > 料理_食材

ご飯をおいしく炊く方法を調べてみると、ざっと以下のようなものが見つか...
見つかったのですが、実際に試してみたことのあるかたの感想が聞きたいです。どれがオススメですか? ご飯をおいしく炊けると言われているウラジ、もち米をひとつかみ加える(米の量の1割程度) ...
解決済み - 更新日時:2009/04/24 03:03:10 - 回答数: 2 - 閲覧数: 344
暮らしと生活ガイド > 料理_グルメ、レシピ

ご飯を1号炊く方法。5、5釜では厳しいと思い、レンジで炊けますか?
5.5釜なら最少1合から炊飯器でいけますよ。あたまも一人暮らしのときに5.5合炊きの釜で1合炊いてました。レンジで炊くならレンジで炊飯できる専用タッパーを使う方法があります。もしくはレンジで使

2. Our system retrieves the top 10
pages of the search results

ご飯を美味しく炊く方法教えてください、条件...

シェア ツイート プクマ

tetsu8191105さん 2012/3/20 15:18:04
ご飯を美味しく炊く方法教えてください、条件電気炊飯器です。
閲覧数: 83 回答数: 1 お礼: 25枚

ベストアンサーに選ばれた回答
mooshing_oleoさん Q&A
米を洗ったあと、しばらくザルにあげる。そのあと水を吸わせる。夏なら最低30分、冬なら最低60分。つけていた水を捨て、別の冷たい水を分量通りいれて、炊く。つけていた水と違う水で炊くと美味しくなるといわれています。炊くときの水は冷たいほうがいいです。
圧力鍋で炊くのが一番美味しいですけどね。

回答詳細 (0件)
ちよいち質問を返す

次に 一目水切りをして30分位置く事

Extracting Tasks

1. Our system performs dependency parsing for all the extracted sentences in order to extract a pattern

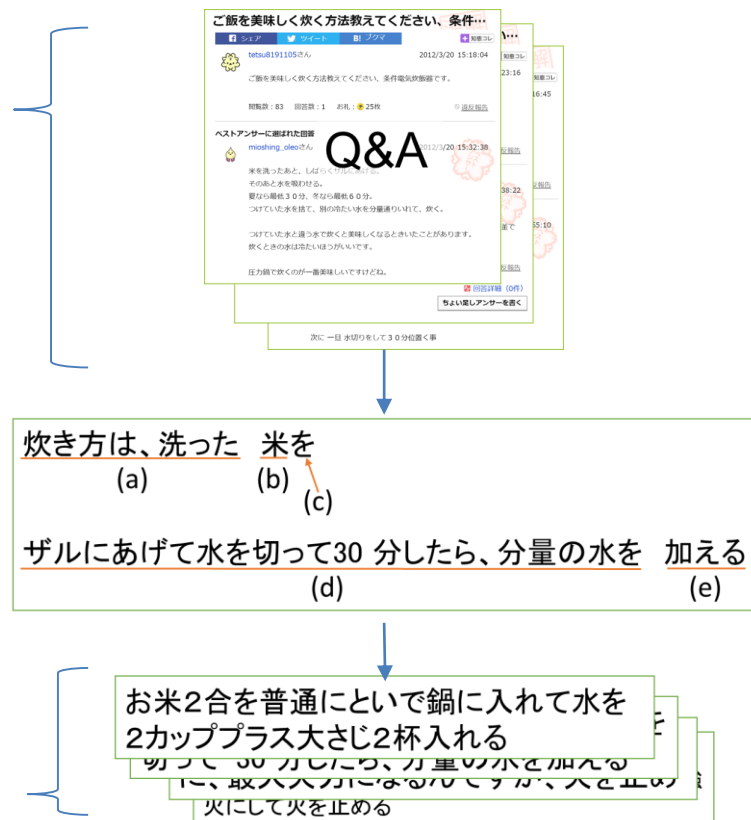
Extracted verb end of the sentence are converted into its dictionary form

2. We extract chunks that contain the following pattern:

Chunks that depend on noun(a) + noun(b) + “を(wo)”(c) + the chunk that includes certain verb(d) + chunks that depend on the verb(e)

3. Our system extracts sentences as tasks

※Considering “を(wo)” is a Japanese particle pointing to a direct object, it is useful to extract tasks that can solve the user's problem



Ranking(1)

- We consider the words that appeared frequently are important in solving problems
- We define the evaluation scores as follows:
- This score consists of NounFreq and VerbFreq

$$\begin{aligned} Score(t_i) = & \sum_{noun \subseteq nouns_{t_i}} \sum_{ans \subseteq A} NounFreq(noun, ans) \\ & + \sum_{task \subseteq T} VerbFreq(verb_{t_i}, task) \end{aligned}$$

→Next I will explain this expression of NounFreq

Ranking (2) Frequency score

“米”(4)+ “ザル”(1)+ “水”(4)+
“分量”(1)+ “水”(4) =14

| Candidate tasks | nouns | frequency |
|--|-------|-----------|
| 炊き方は、洗った米を“ザルにあげて水を切って”30分したら、分量の水を加える | 14 | 4 |
| 鍋を中火にかけて沸騰すれば1~2分間キープして直ぐに火を最低限まで絞って15分間、最後に一瞬強火にして火を止める | 17 | 4 |
| お米2合を普通にといで鍋に入れて水を2カッププラス大さじ2杯入れる | 12 | 3 |
| 言われるように、鍋に米を入れて指の第一関節や手首までお水を入れる | 14 | 3 |
| 米を炊く時は、基本的に吸水させたあとに、最大火力になるんですが、火を止める | 8 | 1 |

$$\sum_{noun \subseteq nouns_{t_i}} \sum_{ans \subseteq A} NounFreq(noun, ans)$$

→Next I will explain this expression of VerbFreq

Ranking(3): VerbFrequency score

| Candidate tasks | verbs | frequency |
|--|-------|-----------|
| 炊き方は、洗った米を“ザルにあげて水を切って”30分したら、分量の水を 加える | 1 | 1 |
| | | 2 |
| 鍋を中火にかけて沸騰すれば1~2分間キープして直ぐに火を最低限まで絞って15分間、最後に一瞬強火にして火を止める | 2 | 2 |
| お米2合を普通にといで鍋に入れて水を2カッププラス大さじ2杯入れる | 2 | |
| 言われるように、鍋に米を入れて指の第一関節や手首までお水を入れる | 2 | |
| 米を炊く時は、基本的に吸水させたあとに、最大火力になるんですが、火を止める | 2 | |

“加える”(1) = 1

$$\sum_{task \subseteq T} VerbFreq(verb_{t_i}, task)$$

Ranking(4):Consists

| Candidate tasks | NounFreq | VerbFreq | Score |
|--|----------|----------|-------|
| 炊き方は、洗った米を“ザルにあげて水を切って”30分したら、分量の水を加える | 14 | 1 | |
| 鍋を中火にかけて沸騰すれば1～2分間キープして直ぐに火を最低限まで絞って15分間、最後に一瞬強火にして火を止める | 17 | 2 | |
| お米2合を普通にといで鍋に入れて水を2カッププラス大さじ2杯入れる | 12 | 2 | |
| 言われるように、鍋に米を入れて指の第一関節や手首までお水を入れる | 14 | 2 | |
| 米を炊く時は、基本的に吸水させたあとに、最大火力になるんですが、火を止める | 8 | 2 | |

$$\begin{aligned} Score(t_i) = & \sum_{noun \subseteq nouns_{t_i}} \sum_{ans \subseteq A} NounFreq(noun, ans) \\ & + \sum_{task \subseteq T} VerbFreq(verb_{t_i}, task) \end{aligned}$$

Ranking(4):Consists

| Candidate tasks | NounFreq | VerbFreq | Score |
|--|----------|----------|-------|
| 炊き方は、洗った米を“ザルにあげて水を切って”30分したら、分量の水を加える | 14 | 1 | 15 |
| 鍋を中火にかけて沸騰すれば1～2分間キープして直ぐに火を最低限まで絞って15分間、最後に一瞬強火にして火を止める | 17 | 2 | 19 |
| お米2合を普通にといで鍋に入れて水を2カッププラス大さじ2杯入れる | 12 | 2 | 14 |
| 言われるように、鍋に米を入れて指の第一関節や手首までお水を入れる | 14 | 2 | 16 |
| 米を炊く時は、基本的に吸水させたあとに、最大火力になるんですが、火を止める | 8 | 2 | 10 |

$$\begin{aligned} Score(t_i) = & \sum_{noun \subseteq nouns_{t_i}} \sum_{ans \subseteq A} NounFreq(noun, ans) \\ & + \sum_{task \subseteq T} VerbFreq(verb_{t_i}, task) \end{aligned}$$

Ranking(5):Score

| Rank | Candidate tasks | NounFreq | VerbFreq | Score |
|------|--|----------|----------|-------|
| 1 | 鍋を中火にかけて沸騰すれば1～2分間キープして直ぐに火を最低限まで絞って15分間、最後に一瞬強火にして火を止める | 17 | 2 | 19 |
| 2 | 言われるように、鍋に米を入れて指の第一関節や手首までお水を入れる | 14 | 2 | 16 |
| 1 | 炊き方は、洗った米を“ザルにあげて水を切って”30分したら、分量の水を加える | 14 | 1 | 15 |
| 4 | お米2合を普通にといで鍋に入れて水を2カッププラス大きじ2杯入れる | 12 | 2 | 14 |
| 5 | 米を炊く時は、基本的に吸水させたあとに、最大火力になるんですが、火を止める | 8 | 2 | 10 |

Our system outputs tasks in the descending order of the score
→The top task includes some frequently appearing words such as “火” and “分”

Experiment

- We mined tasks for the query set of the TaskMine subtask
- Dataset is 50 tasks

Results

- nDCG is computed for a variety of cutoff thresholds $k(=1,5,10,50)$
 - InteractiveMediaMINE performs better than the organizer's baseline method among all the runs

| | nDCG@1 | nDCG@5 | nDCG@10 | nDCG@50 |
|----------------------|--------|--------|---------|---------|
| InteractiveMediaMINE | 0.323 | 0.330 | 0.320 | 0.289 |
| Organizer | 0.013 | 0.040 | 0.053 | 0.096 |

Results(2)

- Our system makes it easy to collect answers accurately for ordinary questions
 - TM-019, 歯周病を治療する(※), has one of the highest scores for all metrics, but extracted task is long sentence as table below

※ “I want to treat periodontal disease”, in English

| Extracted Task | Matched gold standard task |
|--|----------------------------|
| 歯周病治療は、歯周病の原因となる汚れを除去した後、毎日の歯磨きを妨げる歯周ポケットを除去する | 炎症を引き起こす細菌を徹底的に除去する |

- The accuracy of the extracted tasks decreases for queries that are not the type of questions that ordinary users ask
 - TM-023, レーザーカッターを使う(※), has one of the lowest scores for all metrics
- ※ “I want to use the laser cutter”, in English
- Yahoo! Chiebukuro is a general Q & A service that is not specialized in any specific field

Conclusion

- We used Yahoo! Chiebukuro as our system resource because we believe that a Q&A service has many answers that can solve a user's problems
- We experimentally confirmed that a Q&A service is an effective resource for extracting tasks

Future work

- We will add some Q&A services and Internet bulletin boards that are specialized in some specific fields to the proposed system's information resource.
- We plan to improve the ranking method and the precision of the extracting task