Learning Actions Preconditions from Step-by-step instructions in Planning Domains

Maurício Steinert
School of Technology
Pontific Catholic University of Rio Grande do Sul – PUCRS
Porto Alegre, Brazil
Mauricio.steinert@edu.pucrs.br

July-2020

Outline

- Motivation
- Objectives
- WikiHow Dataset
- Embedding dictionaries
- Experiments
  - Actions Identification task
  - Steps preconditions evaluation
- Conclusions

Motivation

- Using unstructured raw data do describe planning domains
  - Images (LatPlan)
  - Text (Framer, StoryFramer)
- In this work, we are interested in using text input

Objectives

- Identify actions within sentences based on word embeddings similarity
- Evaluate if a set of step-by-step instructions respect steps preconditions

WikiHow Dataset

WikiHow Dataset

- Need for a large source of step-by-step instruction on performing tasks
  - WikiHow – https://www.wikihow.org/
- WikiHow raw data structure
  - Step-by-step instructions
  - First sentence of each step summarizes the whole step
  - More complex tasks are divided into sections with independent steps

WikiHow Dataset

- Dataset instance after preprocessing

fill a large pot about 2/3 full of water.
cover the pot and bring the water to a boil.
add salt and 1 pound (450 g) of pasta to the boiling water.
set a timer for 3 to 8 minutes.
stir the noodles occasionally as they boil.
bite into a noodle to see if it’s cooked enough for you.
scoop out about 1 cup (240 ml) of pasta water and set it aside.
set a colander in the sink and put on oven mitts.
pour the pasta into the colander and shake it.
avoid adding oil or running cold water over the pasta if you plan on using sauce.
put the pasta back into the pot and toss it with your choice of sauce.
toss short noodles with pesto or vegetables.
mix cheese into macaroni or shells to make a creamy pasta.
serve meaty sauce over tubular or wide pasta.

WikiHow Dataset

- Dataset instance after preprocessing

1. fill a large pot about 2/3 full of water.

2. cover the pot and bring the water to a boil.

3. add salt and 1 pound (450 g) of pasta to the boiling water.

4. set a timer for 3 to 8 minutes.

WikiHow Dataset

- Dataset instance after preprocessing

1. fill a large pot about 2/3 full of water.

2. cover the pot and bring the water to a boil. Depends on step 1

3. add salt and 1 pound (450 g) of pasta to the boiling water.

4. set a timer for 3 to 8 minutes.

WikiHow Dataset

- Dataset instance after preprocessing

1. fill a large pot about 2/3 full of water.

2. cover the pot and bring the water to a boil. Depends on step 1

3. add salt and 1 pound (450 g) of pasta to the boiling water. Depends on step 2

4. set a timer for 3 to 8 minutes.

WikiHow Dataset

- Dataset instance after preprocessing

1. fill a large pot about 2/3 full of water.

2. cover the pot and bring the water to a boil. Depends on step 1

3. add salt and 1 pound (450 g) of pasta to the boiling water. Depends on step 2

4. set a timer for 3 to 8 minutes. Independent step

WikiHow Dataset

- Statistics
  - 5,519 instances (recipes)
  - 13.75 steps per instance (average)
  - 8.74 words per step (average)

TOP 10 verbs frequency

Verb	Frequency
add	8,305
serve	3,024
place	2,976
remove	2,650
cook	2,490
make	2,292
pour	2,257
cut	2,186
mix	2,168
stir	2,128

Embedding Dictionaries

Embedding dictionaries

- Vector of real numbers represent words
- Captures semantic relationship between words
- Unsupervised learning method
- In this work we use
  - *word2vec* for words embedding
  - *sent2vec* for sentences embedding
  - Trained over Wikipedia raw text
  - 100-dimensional vectors
  - Vocabulary length of approximately 3,000,000 words

Actions Identification task

Actions identification task

- Verbs usually define actions
- Identify verbs within sentences
  - Word embeddings similarity
  - Keywords set
  - Similarity threshold
- Why not using Named Entity Recognition (NER) libraries?
  - We want to evaluate a totally unsupervised method
  - We use them to generate ground-truth values for performance evaluation

Actions identification task

- Experiment setup
  - Keywords: `take, put, drop, give, walk, move, pick, jump, eat, wait, shake, pull, push`
  - Similarity threshold values: 0.3, 0.4, 0.5, 0.6

Actions identification task

- Results:
  - precision, recall, f1-score

Similarity threshold	Precision	Recall	F1-Score
0.3	0.2734	**0.8719**	0.3931
0.4	**0.4958**	0.5449	**0.4951**
0.5	0.2628	0.1944	0.2132
0.6	0.0571	0.0378	0.0433

Actions identification task

- Results
  - Does not perform as we expected
  - Training embedding dictionary in domain content may improve performance
    - Declarative form x imperative form

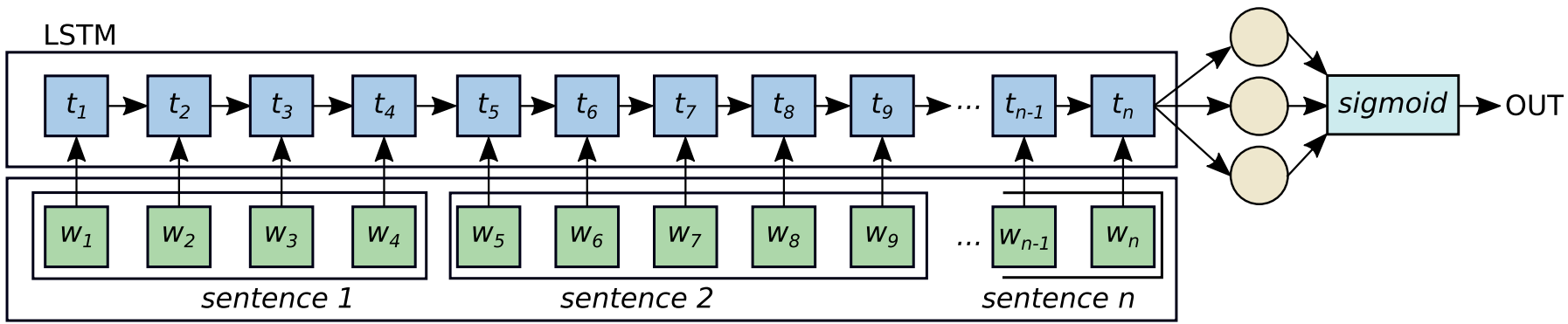
Steps preconditions evaluation

Steps preconditions evaluation

- Evaluate if a set of step-by-step instructions respect steps preconditions
- Classification problem: need for labeled data
  - Normal instance labeled with 1
  - Reverse steps instance labeled with 0
- Train a Recurrent Neural Network with attention
  - Input: embedding representation of text
  - Output: 0 (False) or 1 (True)
  - Input limited to 2,000 values to avoid vanishing gradient

Steps preconditions evaluation

- Experiment setup
  - Training set: 8,831 instances
  - Test set: 2,207 instances
- Recurrent Neural Network architecture



Steps preconditions evaluation

- Results
  - precision, recall, f1-score

Experiment	Precision	Recall	F1-Score
steps-eval-word2vec	0.8615	**0.8315**	0.8498
steps-eval-sent2vec	**0.9235**	0.8312	**0.8749**

Steps preconditions evaluation

- Word2vec - input size limit issue

Normal instance – label 1

fill a large pot about 2/3 full of water.
cover the pot and bring the water to a boil.
add salt and 1 pound (450 g) of pasta to the boiling water.
set a timer for 3 to 8 minutes.
stir the noodles occasionally as they boil.
bite into a noodle to see if it's cooked enough for you.
scoop out about 1 cup (240 ml) of pasta water and set it aside.
set a colander in the sink and put on oven mitts.
pour the pasta into the colander and shake it.
avoid adding oil or running cold water over the pasta if you plan on using sauce.
put the pasta back into the pot and toss it with your choice of sauce.
toss short noodles with pesto or vegetables.
mix cheese into macaroni or shells to make a creamy pasta.
serve meaty sauce over tubular or wide pasta.

Reversed instance – label 0

serve meaty sauce over tubular or wide pasta.
mix cheese into macaroni or shells to make a creamy pasta.
toss short noodles with pesto or vegetables.
put the pasta back into the pot and toss it with your choice of sauce.
avoid adding oil or running cold water over the pasta if you plan on using sauce.
pour the pasta into the colander and shake it.
set a colander in the sink and put on oven mitts.
scoop out about 1 cup (240 ml) of pasta water and set it aside.
bite into a noodle to see if it's cooked enough for you.
stir the noodles occasionally as they boil.
set a timer for 3 to 8 minutes.
add salt and 1 pound (450 g) of pasta to the boiling water.
cover the pot and bring the water to a boil.
fill a large pot about 2/3 full of water.

Missing information

Steps preconditions evaluation

- Results
  - Both experiments are capable of learning relationship between steps
  - *Steps-eval-word2vec*
    - Information loss of approximately 76% due to input truncated data
  - *Steps-eval-sent2vec*
    - Performed better than *steps-eval-word2vec*
    - Receives a compact but full of whole task

Conclusions

- Actions identification task based on word similarity does not perform as expected
- Learning algorithms are able to identify relationship between steps
- Future research
  - Overcome vanishing gradient issues: try other NN architectures (Transformer)
  - Identify specific preconditions for each step
  - Perform task hierarchical organization (HTN)