Representing Verbs as Argument Concepts

Yu Gong, Kaiqi Zhao and **Kenny Q. Zhu**
Shanghai Jiao Tong University
Dec 27, 2015

Outline

- Introduction
- Related Work
- Problem Definition
- Approach
- Experiments
- Conclusion

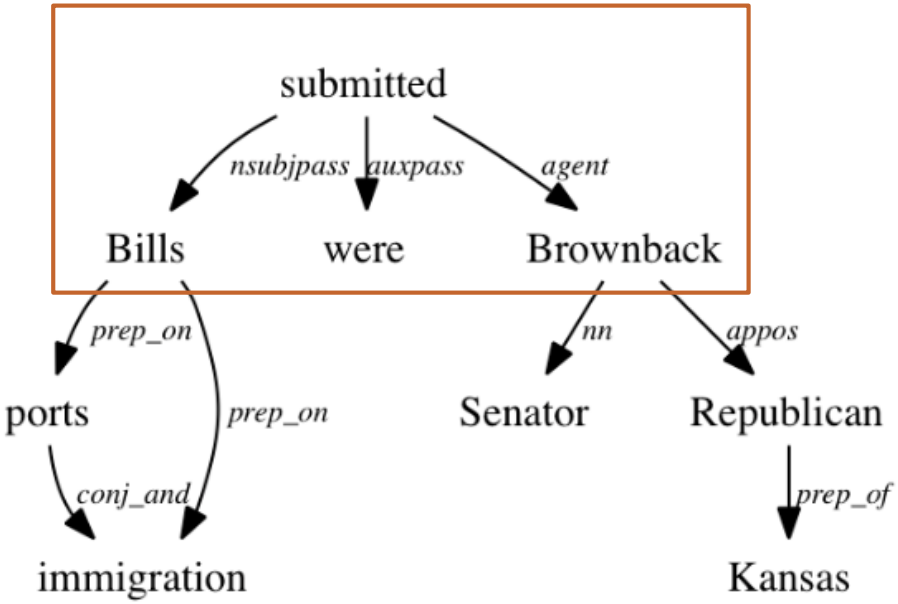
Introduction

- Representation of a verb
  - It is possible to represent the meaning of a word by the distributional properties of its context.
    - Word2Vec
  - A verb is unique in a sentence that it maintains dependency relation with its syntactic arguments such as the subject and the object.

Introduction

Bills on ports and immigration were submitted by Senator Brownback, Republican of Kansas

window window

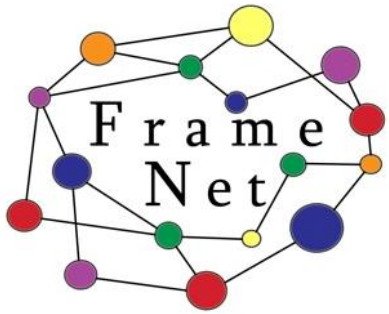


Introduction

- Why argument concepts?
  - Possible to use the distribution of immediate arguments of a verb to represent its meaning.
    - The naïve method is “Bag of Words” (BoW)
    - BoW method has many limitations
      - Independence between words
      - High dimensionality
      - Poor readability
    - So, we represent the arguments by their abstract types

Related Work

- Semantic Role Labeling (SRL)
  - Use a lexicon to define the semantic roles of the arguments of that verb.
    - e.g. FrameNet, PropBank or ReVerb
    - Eat → Ingestion→Ingestibles



Related Work

- Semantic Role Labeling (SRL)
  - Limitations:
    1. Human annotation is required, which limits their scales.
    2. The frames are course-grained: unable to distinguish between two close senses.
    3. Semantic roles in SRL are used as labels only: no relationships among the labels; not computable

Related Work

- ReVerb
  - It is an open information extraction system to discovers verb triples from web.
  - It is too fine-grained.
  - It is lack of abstraction:
    - a system powered by ReVerb will not recognize a verb and its arguments unless ReVerb has this triple in the knowledge.

Related Work

- Selectional Preference (SP)
  - With a taxonomy, SP can produce a ranked list of concepts that are the most appropriate subjects or objects of a predicate verb.
  - The definition of selectional association:

$$A(p,c) = \frac{\Pr(c|p)\, log \frac{Pr(c|p)}{Pr(c)}}{\sum_{c' \in C} Pr\,(c'|p) log \frac{Pr(c'|p)}{Pr(c')}}$$

Related Work

- Selectional Preference (SP)
  - The limitations:
    1. Not consider the diversity of concepts, which may give a list of concepts with the same meaning.
    2. It assumes every argument to the verb is correct and contributes to the selectional strength, but action instances obtained by parsing are often noisy and contain errors.

Problem Definition

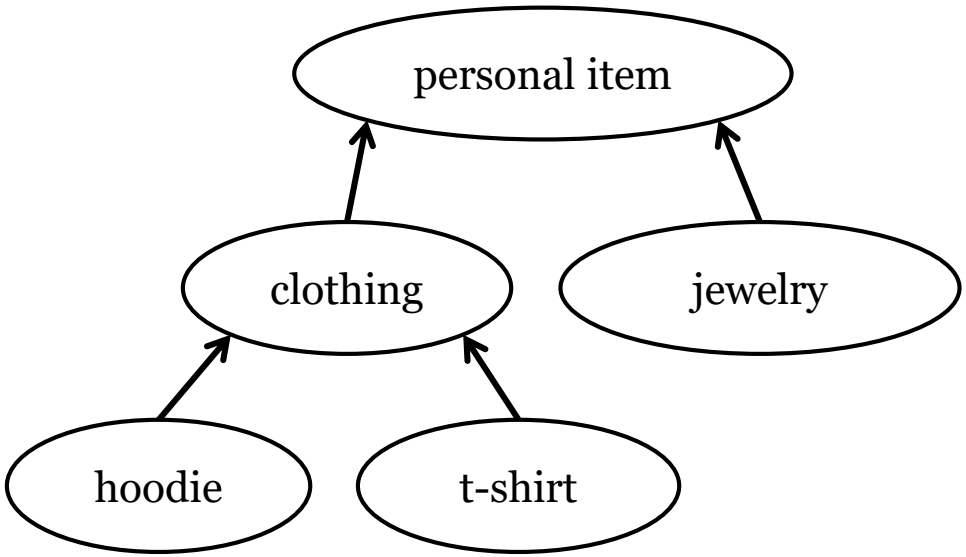
- Informal Definition
  I. Given a collection of argument instances (either subjects or objects) for a verb;
  II. Pick *k* concepts from the taxonomy that subsume as many instances as possible, which is equivalent to maximizing the likelihood of the corpus.
  III. We would like these *k* concepts to have little overlap against each other.

Problem Definition

- Informal Definition
  - Intuition
    - Each of the *k* selected concepts represents a unique semantic and the *k* concepts collectively cover majority of the uses of that verb.
  - Example
    - Argument Instances:
      - wear/{t-shirt, hoodie, hat, bracelet, ear ring, pink}
    - Argument Concepts:
      - wear/{clothing, accessory, style}

Problem Definition

- Taxonomy



Problem Definition

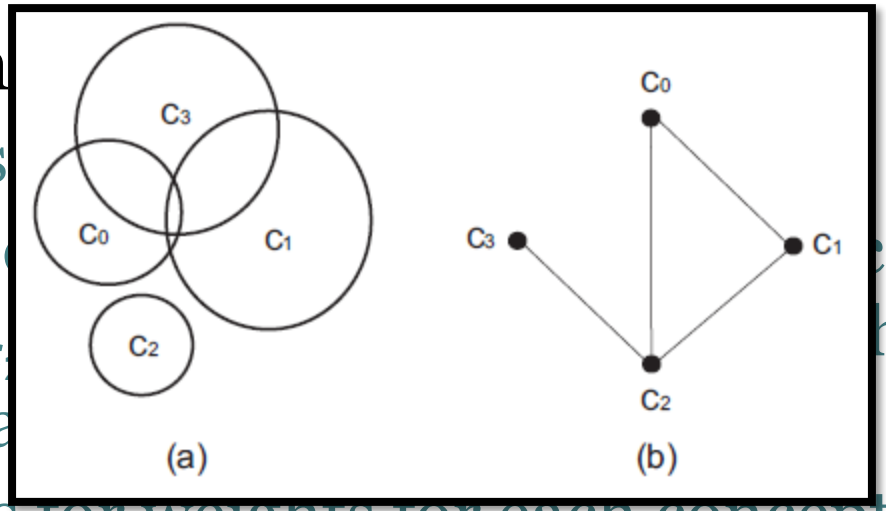
- Definition 1. *Overlap*:
  - The overlap between two concepts is

$$Overlap(c_1, c_2) = \frac{|E_{c_1} \cap E_{c_2}|}{min\{|E_{c_1}|, |E_{c_2}|\}}$$

where E_c is the set of all entities covered by concept c in the taxonomy.

Problem Definition

- Definition (C, L, W) :
  - C is the s
  - L is the s epts. An edge $l_{c_1,c}$ hen overlap lower tha
  - W stands for weights for each concepts in the graph, which represents the quality of the concept with respect to the verb.



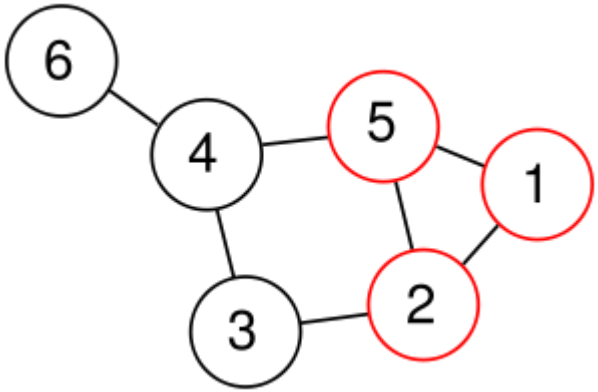
Problem Definition

- Definition 3. Concept Weight $w_v(c)$:
  - The naïve method is counting the number of argument instances it subsumes according to the isA taxonomy (baseline).
  - But all argument instances of a verb are not of equal importance, so we define Quality Function $Q_v(e)$

$$w_v(c) = \sum_{e \in \{e | e \; isA \; c\}} Q_v(e)$$

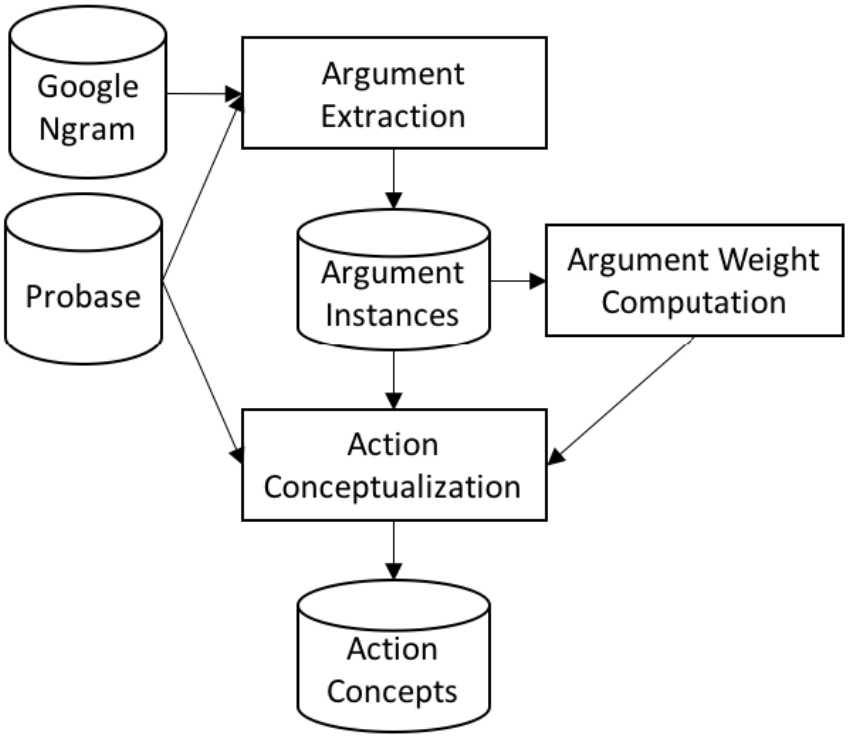
Problem Definition

- Definition 4. *Argument Conceptualization*:
  - The problem is transformed to finding the *k*-clique with maximum combined weight.
  - It is proved to be *NP-Complete*.



Approach

- System Overview



Approach

- Argument Weight Computation
  - *Entropy*
    - Dependency Parser may lead to errors.
    - But, some errors follow certain patterns.

"food to eat" "water to drink" "game to play"

"play this time"

Approach

- Argument Weight Computation
  - *Entropy*
    - If an argument is incorrect due to parsing, it is often extracted from just a *few* patterns.
    - Conversely, if an argument is correct for the verb, it should appear under *different* patterns.

"eat meat" "eat expensive meat"

"eat not only meat"

Approach

- Argument Weight Computation
  - *Entropy*
    - We define a pattern as a subtree in the dependency tree according to the following rules:
      - The argument and one of its child:

        $\{POS_{arg}, DEP_{arg}, POS_{child}, DEP_{child}\}$
      - The argument and its sibling:

        $\{POS_{arg}, DEP_{arg}, POS_{sib}, DEP_{sib}\}$

Approach

- Argument Weight Computation
  - *Entropy*
    - For each argument e of verb v, we collect the set of its patterns $M_{e,v}$, and an argument that appears in more patterns has higher probability to be correct, and thus has higher quality.
    - We use the entropy to measure the correctness:

$$Entropy_v(e) = -\sum_{m \in M_{e,v}} P(m) log P(m)$$

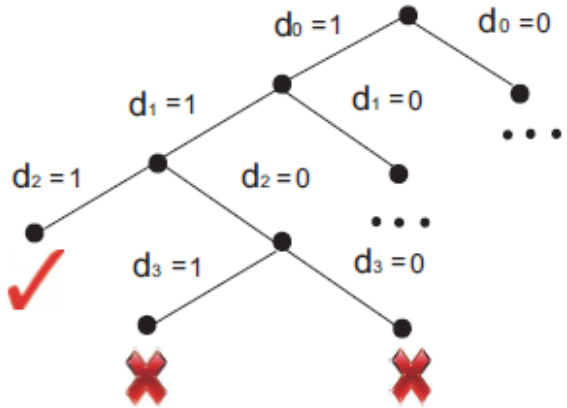
Approach

- Argument Weight Computation
  - *Mutual Information*
    - A measure can capture the strength of mutual connection between two terms.
    - E.g. "eat thing" and "eat fruit"
    - We use binary version of MI

$$MI_v(e) = \begin{cases} 1 & if\ p(v,e)log\dfrac{p(v,e)}{p(v)} > 0, \\ -1 & otherwise \end{cases}$$

Approach

- A Branch-and-Bound Algorithm
  - *The Branch*
    - Generate a decision tree
    - The nodes at each level represent the decision to include a concept in the solution or not.
    - A path in the tree is a candidate solution.



Approach

- A Branch-and-Bound Algorithm
  - *The Bound*
    I. ISCLIQUE
       - The current path must be a clique with the size no lager than k.

    II. BOUND
       - Maximum possible score is larger than current best score.

Experiment

- Experiment Setup
  - IsA taxonomy

ProBase

WordNet
A lexical database for English

  - The dataset

Google N-grams

Experiment

- Conceptualization Results
  - three English speakers to annotate whether the concepts generated by AC, BL and SP are the correct abstraction of the verb's arguments

k	Subject			Object		
	AC	BL	SP	AC	BL	SP
5	**0.88**	0.49	0.58	**0.97**	0.63	0.62
10	**0.86**	0.47	0.56	**0.94**	0.61	0.65
15	**0.85**	0.43	0.58	**0.91**	0.60	0.66

Table 1: Example subject/sbject concepts from 4 lexicons

Verb		Action Concepts	FrameNet	ReVerb	SP Concepts
accept	Subj	person,community, institution,player,company	Recipient,Speaker, Interlocutor	Student,an article, the paper,Web browser, Applications	world,target group, group,term,person
	Obj	document,payment, practice,doctrine,theory	Theme,Proposal	the program,publication, HTTP cookie,the year, credit card	topic,concept,matter, abstract entity,document
cause	Subj	factor,disease, event,agent,technique	Actor	The root,HIV, Car accident,Suicide, Cardiovascular disease	word,factor,condition complication,symptom
	Obj	disease,effect, challenge,emergency,defect	Event	the problem,AIDS, Poverty,death, Heath problems	symptom,complication, condition,disease,factor
consume	Subj	factor,product, person,feature,activity	Ingestor	people,a person, fire,The products, United States	world,company, characteristic,factor,term
	Obj	food,substance,industry species,product and service	Ingestibles	information,Sacrifice, news,Alcoholic beverage, the burnt offering	unit,information, food,number,term
enjoy	Subj	group,community, name,country,sector	Experiencer	people,ive,Guests, everyone,someone	person,group,world,actor vulnerable population
	Obj	benefit,time,hobby, social event,attraction	Stimulus	life,Blog,Breakfirst, their weekend,a drink	benefit,issue, advantage,topic,quality
plan	Subj	name,group,topic, community,item	Agent	God,master,couples, Work,action	world,name,person, group,company
	Obj	service,event,factor, place,organization	Goal,Event	our lives,communities, all,Wedding,FY 2001	event,activity,area, project,word

Experiment

- Argument Identification
  - use the inferred argument concepts to examine whether a term is a correct argument to a verb in a sentence

	k	Probase			WordNet			RV \ SRL
		AC	BL	SP	AC	BL	SP	
Subj	5	**0.81**	0.50	0.70	0.55	0.54	0.54	0.48 / 0.54
	10	**0.78**	0.50	0.72	0.57	0.54	0.55	
	15	**0.77**	0.49	0.72	0.58	0.54	0.56	
Obj	5	**0.62**	0.51	0.58	0.50	0.46	0.50	0.50 / 0.47
	10	**0.62**	0.52	0.58	0.52	0.47	0.52	
	15	**0.62**	0.52	0.59	0.53	0.47	0.52	

Action Conceptualization

Ex

- D

Input the verb: [ ] Submit

A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

abate
abduct
abolish
abrogate
abuse
accede
accelerate
accentuate
accept
acclaim
accommodate
accomplish
accrue
accumulate
acetylate
ache
achieve
acknowledge
acquiesce
acquire
acquit
act
activate
adapt
add
addict
address

accept k = 5

Subject

Action Concepts	SP Concepts
person	world
community	target group
institution	group
player	term
company	person

Object

Action Concepts	SP Concepts
document	topic
payment	concept
practice	matter
doctrine	abstract rntity
theory	document

Conclusion

- Argument instances parsed from raw text
- Abstract into concepts that is:
  - Human readable
  - Machine computable
  - Representation of the verb
- Shows good results in argument identification
- More NLP tasks such as WSD, similarity…

Thanks!