

Relation Extraction (RE) via Supervised Classification

See:

- Jurafsky & Martin SLP book, Chapter 22
- Exploring Various Knowledge in Relation Extraction. ZHOU GuoDong SU Jian ZHANG Jie ZHANG Min, ACL 2005.

Relations between Entities

- Classification instance: a (ordered) pair of entities
 - Typically in a sentence
 - Arguments not always named entities, can be common noun phrases (e.g. for *attack*)
 - This requires segmentation (IOB – like NER)
- May target single or multiple relations
- Annotated training for relation instances
 - relation type, argument spans and their roles
 - Negative examples may be all entity pairs that are not annotated as having a relation
- A restricted case of *Information Extraction (IE)*
 - Extract richer templates

Classification Architectures

- Binary class for each relation, one-versus-all
 - Highest classification score wins (or ranking of positives)
 - All classifications negative implies no relation
- Multi-class classifiers, with no-relation as a class
- Two tier classification:
 - Is there a relation? (binary)
 - Relation type
 - multi-class, possibly one vs. all (highest negative score may win)
- Argument role may be distinguished by its NER type (e.g. employee-of), or by directional features
- May classify each participant to its role (vs order)
 - Usually done in template-filling IE

Features

(based on James Martin – 4 slides)

- We can group the RE features into three categories
 - Features of the named entities/arguments involved
 - Features derived from the words between and around the named entities
 - Features derived from the syntactic environment that governs the two entities

Features

- Features of the entities
 - Their types
 - Concatenation of the types
 - Headwords of the entities
 - *George Washington Bridge*
 - Words in the entities
 - Notice: arguments aren't only named entities, can be (common-) noun phrases
- Features between and around
 - Particular positions to the left and right of the entities
 - +/- 1, 2, 3
 - Bag of words/ n-grams between
 - Words related to the predicate words, e.g. WordNet synonyms

Features

- Syntactic environment
 - Constituent path through the tree from one to the other
 - Base syntactic chunk sequence from one to the other
 - Dependency path
 - Indicators of certain edges/labels along the path
 - E.g. apposition
 - Tree-distance between arguments

Example

- For the following example, we're interested in the possible relation between American Airlines and Tim Wagner.
 - *American Airlines*, a unit of AMR, immediately matched the move, spokesman *Tim Wagner* said.

Entity-based features

Entity ₁ type	ORG
Entity ₁ head	<i>airlines</i>
Entity ₂ type	PERS
Entity ₂ head	<i>Wagner</i>
Concatenated types	ORGPERS

Word-based features

Between-entity bag of words	{ <i>a, unit, of, AMR, Inc., immediately, matched, the, move, spokesman</i> }
Word(s) before Entity ₁	NONE
Word(s) after Entity ₂	<i>said</i>

Syntactic features

Constituent path	$NP \uparrow NP \uparrow S \uparrow S \downarrow NP$
Base syntactic chunk path	$NP \rightarrow NP \rightarrow PP \rightarrow NP \rightarrow VP \rightarrow NP \rightarrow NP$
Typed-dependency path	$Airlines \leftarrow_{subj} matched \leftarrow_{comp} said \rightarrow_{subj} Wagner$

Tuning and Analysis

- *Look at the data*
- Examine feature weights
 - most positive/negative
- Analyze classification errors
 - False positives, false negatives
- Try alternative feature selection policies

What about lexical variability?

- Relevant for both relation and argument words
- Without external resources - variability needs to be covered in training data
- External lexical similarity resources, manual and/or statistical, may be used for “lexical expansion”; but it’s not trivial to gain substantial benefit from them in a supervised setting
 - WordNet, WikiData, ... vs. PPDB, DIRT, embeddings...
 - DIRT-style rules may be useful for relation variability, there has been work in this direction in the IE field

Template/Event Information Extraction

- Goal: extract complete templates with slots, often about events
 - attack, acquisition, conviction, ...
- Extending the RE supervised scheme
- Possible architecture
 - Classifier for event trigger
 - Classifier for each slot
 - Possibly joint classification rather than pipeline

WordNet – an Online Lexical Database

<http://wordnet.princeton.edu/>

- High coverage lexical-semantic network built by psychologists

POS	Monosemous Words and Senses	Polysemous Words	Polysemous Senses
Noun	101863	15935	44449
Verb	6277	5252	18770
Adjective	16503	4976	14399
Adverb	3748	733	1832
Totals	128391	26896	79450

- Relations:
 - ISA-relation (hyponom - hypernym, taxonomic backbone)
 - Part-of (meronym - holonym)
 - Type-instance (e.g. Obama is an instance of President)
 - Opposite-of (antonym), mostly for adjectives
 - Derivative (pertainym), e.g. crime – criminal
 - some semantic roles between verbs and nouns, e.g. AGENT, INSTRUMENT ...
- A useful source for lexical expansions/inferences (but coverage limited)

Synsets for “magazine#n”

WordNet Search - 3.0 - [WordNet home page](#) - [Glossary](#) - [Help](#)

Word to search for:

Display Options:

Key: "S:" = Show Synset (semantic) relations, "W:" = Show Word (lexical) relations

Noun

Synset

sample use

- (13)**S:** (n) **magazine#1**, [mag#1](#) (a periodic publication containing pictures and stories and articles of interest to those who purchase it or subscribe to it) *"it takes several years before a magazine starts to break even or make money"*
- (2)**S:** (n) **magazine#2** (product consisting of a paperback periodic publication as a physical object) *"tripped over a pile of magazines"*
- (1)**S:** (n) **magazine#3**, [magazine publisher#1](#) (a business firm that publishes magazines) *"he worked for a magazine"*
- **S:** (n) **magazine#4**, [cartridge#2](#) (a light-tight supply chamber holding the film and supplying it for exposure as required)
- **S:** (n) **magazine#5**, [powder store#1](#), [powder magazine#1](#) (a storehouse (as a compartment on a warship) where weapons and ammunition are stored)
- **S:** (n) [cartridge holder#1](#), [cartridge clip#1](#), [clip#1](#), **magazine#6** (a metal frame or container holding cartridges; can be inserted into an automatic gun)

gloss

SemCor count

Lexical members

WordNet Hypernym Chain

- (2) S: (n) magazine#2 (product consisting of a paperback periodic publication as a physical object) *"tripped over a pile of magazines"*
 - direct hypernym / inherited hypernym / sister term
 - S: (n) product#2, production#3 (an artifact that has been created by someone or some process) *"they improve their product every year"; "they export most of their agricultural production"*
 - S: (n) creation#2 (an artifact that has been brought into existence by someone)
 - S: (n) artifact#1, artefact#1 (a man-made object taken as a whole)
 - S: (n) whole#2, unit#6 (an assemblage of parts that is regarded as a single entity) *"how big is that part compared to the whole?"; "the team is a unit"*
 - S: (n) object#1, physical object#1 (a tangible and visible entity; an entity that can cast a shadow) *"it was full of rackets, balls and other objects"*
 - S: (n) physical entity#1 (an entity that has physical existence)
 - S: (n) entity#1 (that which is perceived or known or inferred to have its own distinct existence (living or nonliving))

