Intelligent Assistant Agent based on {subject, predicate, object} triplets

Abstract

State of the Art

→ care este actuala tehnologie pentru agentii inteligenti

Motivation and Proposal

→ de ce modelele bazate pe ml si deep learning sunt un “black box” si pentru un agent inteligent conversational este mai utila o abordare clasica, in care putem cunoaste la orice moment state-ul intern al bazei de cunostinte si putem afla de ce agentul a ales un anumit raspuns la o anumita intrebare

Intro

“a triplet in a text sentence is defined as a relation between subject and object, the relation being the predicate.The aim here is to extract sets of the form {subject, predicate, object} out of syntactically parsed sentences.The triple is a minimal representation for information without losing the context.”

POS:

The list is as follows:

* [ADJ](http://universaldependencies.org/u/pos/ADJ.html): adjective
* [ADP](http://universaldependencies.org/u/pos/ADP.html): adposition
* [ADV](http://universaldependencies.org/u/pos/ADV.html): adverb
* [AUX](http://universaldependencies.org/u/pos/AUX_.html): auxiliary verb
* [CONJ](http://universaldependencies.org/u/pos/CONJ.html): coordinating conjunction
* [DET](http://universaldependencies.org/u/pos/DET.html): determiner
* [INTJ](http://universaldependencies.org/u/pos/INTJ.html): interjection
* [NOUN](http://universaldependencies.org/u/pos/NOUN.html): noun
* [NUM](http://universaldependencies.org/u/pos/NUM.html): numeral
* [PART](http://universaldependencies.org/u/pos/PART.html): particle
* [PRON](http://universaldependencies.org/u/pos/PRON.html): pronoun
* [PROPN](http://universaldependencies.org/u/pos/PROPN.html): proper noun
* [PUNCT](http://universaldependencies.org/u/pos/PUNCT.html): punctuation
* [SCONJ](http://universaldependencies.org/u/pos/SCONJ.html): subordinating conjunction
* [SYM](http://universaldependencies.org/u/pos/SYM.html): symbol
* [VERB](http://universaldependencies.org/u/pos/VERB.html): verb
* [X](http://universaldependencies.org/u/pos/X.html): other

Dependences:

ACL: Clausal modifier of noun

* ACOMP: Adjectival complement
* ADVCL: Adverbial clause modifier
* ADVMOD: Adverbial modifier
* AGENT: Agent
* AMOD: Adjectival modifier
* APPOS: Appositional modifier
* ATTR: Attribute
* AUX: Auxiliary
* AUXPASS: Auxiliary (passive)
* CASE: Case marker
* CC: Coordinating conjunction
* CCOMP: Clausal complement
* COMPOUND: Compound modifier
* CONJ: Conjunct
* CSUBJ: Clausal subject
* CSUBJPASS: Clausal subject (passive)
* DATIVE: Dative
* DEP: Unclassified dependent
* DET: Determiner
* DOBJ: Direct Object
* EXPL: Expletive
* INTJ: Interjection
* MARK: Marker
* META: Meta modifier
* NEG: Negation modifier
* NOUNMOD: Modifier of nominal
* NPMOD: Noun phrase as adverbial modifier
* NSUBJ: Nominal subject
* NSUBJPASS: Nominal subject (passive)
* NUMMOD: Number modifier
* OPRD: Object predicate
* PARATAXIS: Parataxis
* PCOMP: Complement of preposition
* POBJ: Object of preposition
* POSS: Possession modifier
* PRECONJ: Pre-correlative conjunction
* PREDET: Pre-determiner
* PREP: Prepositional modifier
* PRT: Particle
* PUNCT: Punctuation
* QUANTMOD: Modifier of quantifier
* RELCL: Relative clause modifier
* ROOT: Root
* XCOMP: Open clausal complement

Universal Dependencies (UD) is a project that is developing cross-linguistically consistent treebank annotation for many languages, with the goal of facilitating multilingual parser development, cross-lingual learning, and parsing research from a language typology perspective. The annotation scheme is based on an evolution of (universal) Stanford dependencies (de Marneffe et al., 2006, 2008, 2014), Google universal part-of-speech tags (Petrov et al., 2012), and the Interset interlingua for morphosyntactic tagsets (Zeman, 2008). The general philosophy is to provide a universal inventory of categories and guidelines to facilitate consistent annotation of similar constructions across languages, while allowing language-specific extensions when necessary.

<https://github.com/NSchrading/intro-spacy-nlp> → subject\_object\_extraction file source (bad)

NOTES:

Every conversation is a context → some facts have sense only in specific contexts

Synonyms in contexts: Mary likes Tom

Does Mary admire Tom? → YES

Calculate word similarities!

**TODO:**

**\*** extract properties !!!

\* get triplets from questions !!!

\* construct rdf graph

\* query existing graph

\* documentation

To look at:

- Cases where there is no subject or object !!

- Subject attributes

- Prepositions (“in shapes” instead of “shapes”)