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

Style Notes:

Use EndNote, apa citation style, scholar.google.ro

Abstract

State of the Art

→ care este actuala tehnologie si abordare pentru agentii inteligenti/ agenti conversationali

\* Focus Google Assitant (Others: Siri, Alexa, Cortana)

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

→ modelul bazat pe triplete si baza de cunostinte se pliaza mai bine pe schimbarile din limbaj (o retea trebuie reantrenata); conexiune limbaj vs lume

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.”

spacy description

Implementare

probleme aparute + cum le-am rezolvat

aborare generala

descriere spacy, rdf, sparql

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

**Spacy** generates the DG ( dependencies Grammar).. I do the rest!

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

→ first search in the actual context graph, and then in the rest

Synonyms in contexts: Mary likes Tom

Does Mary admire Tom? → YES

Calculate word similarities!

**TODO:**

\* generate queries from all types of questions

\* intelligent sewing machine problem

Work in progress:

- queries

Comparison:

Pattern No. Example and Triplets Extracted

S1V1O1

The system displays the metadata to the data manager

(system, displays metadata to, data manager)

(system, displays, metadata)

('system', 'displays', 'metadata'), ('system', 'displays', 'to data manager')

SnV1O1

Either administrator or data manager should be good in database.

(data manager, should be good in, database)

(administrator, should be good in, database)

[('administrator', 'good', 'in database'),

('data manager', 'good', 'in database')]

S1VnO1

Administrators are responsible for installing, configuring and monitoring the system.

(Administrators, configuring, system)

(Administrators, installing, system)

(Administrators, monitoring, system)

[('administrators', 'responsible', 'system')]

S1V1On

Administrators should have skills in system administration, database management and deployment of Web applications.

(administrators, should have, skills),

(administrators, should have skills in, database management),

(administrators, should have skills in, system administration),

(administrators, should have skills in, deployment),

(administrators, should have skills in deployment of, web applications)

[('administrators', 'have', 'in database management'),

('administrators', 'have', 'in deployment'),

('administrators', 'have', 'in system administration'),

('administrators', 'have', 'of web applications'),

('administrators', 'have', 'skills')]

SnV1On

Portal managers, Data managers should have basic knowledge of taxonomy and biodiversity data.

(portal managers, should have basic knowledge of, taxonomy),

(portal managers, should have basic knowledge of, biodiversity data),

(data managers, should have basic knowledge of, taxonomy),

(data managers, should have basic knowledge of, biodiversity data).

SnVnO1

IPT instances and GBIF portal Web services, other data source types may be configured or updated as modules.

(data source types, may be configured as, modules),

(portal web services, may be configured as, modules),

(ipt instances, may be configured as, modules),

(data source types, updated as, modules),

(portal web services, updated as, modules),

(ipt instances, updated as, modules).

[('data managers', 'have', 'knowledge'),

('data managers', 'have', 'of data'),

('knowledge', 'property', 'basic')]

S1VnOn

The data manager wants to manage resources to import a new resource, edit a resource metadata or delete a resource.

(data manager, delete, resource metadata),

(data manager, wants to manage, resources),

(data manager, to import, new resource),

(data manager, delete, resource),

(data manager, edit to manage, resources),

(data manager, edit, resource metadata)

split error

SnVnOn

The data manager and administrator wants to manage jobs so as to monitor finished and upcoming jobs, schedule a new job or cancel a scheduled job.

(administrator, wants to manage, jobs)

(data manager, wants to manage, jobs)

(administrator, to monitor, finished and upcoming jobs)

(data manager, to monitor, finished and upcoming)

(data manager, schedule, new job)

(administrator , schedule, new job)

(administrator, cancel, scheduled job)

(data manager, cancel, scheduled job)

a lot of nulls

SiViOi

user and visitor should be able to conduct a search by providing either restaurant name, restaurant description.

(user , to conduct search by providing, restaurant name),

(user , to conduct search by providing, restaurant description),

(visitor , to conduct search by providing, restaurant name),

(visitor, to conduct search by providing, restaurant description).

bad