

CS 486 Introduction to Artificial Intelligence

Fall 2016

Assignment 4: A Natural Language Generation “Hackathon” – Part (a)

Group names (surname first):

Son, Dongwoo (20420487)

Cho, Hyeiwon (20462909)

Experimental Plan

For this assignment, the automated natural language generation is built and tested using NaturalOWL. NaturalOWL is a natural language generation engine. Given ontology file (.owl), it identifies the classes and individuals, and automatically generates the text that describes it based on its properties (and if it is an individual, the classes it belongs to are also used to describe the individual). Without natural language resources, the sentences are very likely to be grammatically incorrect or do not make much sense. Building up the natural language resources changes how the text is generated. In this assignment, two issues of generating sentences based on the ontology using natural language resources will be examined.

A) Research plan:

- The first issue to research in this assignment is Natural Language Resource issues. Initially, simple Natural Resources file given in the course webpage (NLResources.owl) will be used to generate the text. We will build and update the natural language resource and see how different changes and addition of natural language resource will affect the text generated.
- The other issue to research is Knowledge Representation issues, that is, we will experience how the structure of the ontology itself affects the natural language generation. We will experiment this issue by specifying additional and more detailed classes and add more properties as updating the natural language resources associated with the new classes and properties.

B) Primary Domain Ontology

- The primary domain ontology is about the book (or movie) Harry Potter – contains the classes about different creatures, school features and the individuals are the characters. The Harry Potter example ontology (harrypotter0.owl) provided in this course is used as the base ontology to build up the natural language in this assignment. Also, as mentioned in A) above, NLResource.owl is used as our base natural language resources file to be experienced and updated with.

C) Planned extensions to the domain ontology or resources

- To research about the issue of Natural Language Generation, we plan to create some basic Lexicon entries, NL Names associated with the classes and individuals in the ontology. Also, multiple sentence plans will be developed to give the baseline of sentence structure to the properties.

- To test the Knowledge Representation issue, we plan to add several specialized subclasses and properties to the base ontology, as well as adding up natural language resources to the new classes and properties. We will see how changes and additions to the ontology itself affect the Natural Language Resource generation.
- D) Hypothesis for testing – there are 2 hypothesis for this research
- i. Creating additional sub-classes and properties for a given ontology will lead to more expressive text.
 - ii. Adding more sentence plans and other natural language resources will lead to more fluent text.

Description of Parts of the Plan Actually Implemented

Lexicon Entries

Many lexicon entries are added (mostly nouns) to specify the basic structure and properties (e.g., Gender, tense if verb) of the word. The lexicons are used to generate appropriate word strings when building NL Names and Sentence Plans.

Adjectives:

- dark, interested, magical, old, such

Nouns:

- albus-dumbledore, aurther-weasley, being, classmate, cloak, course, deatheater, draco-malfoy, dudley-dursley, father, fawkes, ginny-weasley, headmaster, hedwig, hermione-granger, Hogwarts, house, housemember, husband, invisibility, james-potter, lily-potter, Lucius-malfoy, magic, molly-weasley, mother, muggle, nagini, non-magical-being, owl, parent, petunia-dursley, phoenix, professor, Ronald-weasley, school, shapeshifter, slytherin-house, snake, Vernon-dursley, wife, werewolf, year

Verbs:

- belongVerb, ownByVerb, teachVerb, marryVerb

NL Names

NL Names are used to specify names of the individuals and classes in the ontology. It uses lexicon entries to generate appropriate form of strings for the character.

English NL Names:

- AlbusDumbledore, ArthusWeasley, DarkWizard, Deatheater, DracoMalfoy, Father, Fawkes, GinnyWeasley, Headmaster, Hedwig, HermioneGranger, Hogwarts, House, HouseMember, JamesPotter, LilyPoter, LuciusMalfoy, MagicalBeing, MollyWeasley, Mother, Muggle, Nagini, NonMagicalBeing, Owl, PetuniaDursley, Phoenix, Professor, RonaldWeasley, SchoolOfMagic, Snake, TheClockOfInvisibe, VernonDursley, Werewolf, Wizard

Sentence Plans

Sentence Plans are used to give the basic structure for generating sentences for properties, and specify the sentence structure for specified properties. It is useful when generating smooth sentence of property object and property fillers.

English Sentence Plans:

- ClassmateOf, DaughterOf, FatherOf, HouseOf, HouseMemberOf, HusbandOf, MotherOf, OwnedBy, ParentOf, hasStudent, isInterestedIn, isNYearsOld, teachesCourses, isMarriedTo, WifeOf

Specialized Classes

This classes are added to experiment the change in sentence generation.

- Gender class, which has subclasses Male and Female.
- Mother and Father classes, which are subclasses of Parent
- Daughter and Son classes, which are subclasses of Child
- Wife and Husband classes, which are subclasses of MarriedCouple

Specialized Properties

More properties are added, more sentence plans associated with these properties are created.

- isFatherOf and isMotherOf properties, which are sub-properties of isParentOf
- isDaughterOf and isSonOf properties, which are sub-properties of isChildOf
- isWifeOf and isHusbandOf properties, which are sub-properties of isMarriedTo
- hasCourses property, which is used as assertion with a domain of SchoolOfMagic and range of Course
- isClassMateOf properties, which is used as assertion with a domain of Student and range of Student
- hasStudent properties, which is used as assertion with a domain of House and range of Student. It is inverse property of isHouseMemberOf

What worked well vs What didn't work well

The research generally worked well, as a result, we were able to generate the fairly smooth, and grammatically correct sentences for most of the texts generated based on the ontology. However, there were some incorrect grammars and repetition of the world and sentences for the individuals under the same classes. The hypothesis are correct - as we build more natural language resources we get more accurate sentences, and as we specify more classes and properties, the result of the text generated was enriched, making more description for the individuals. The result of the research is more detail explained in part d documentation (a4_partD_son-cho.pdf).