Symbolic AI Coursework - The Riddle of Steel

2019.05.07

Haorui_Chen

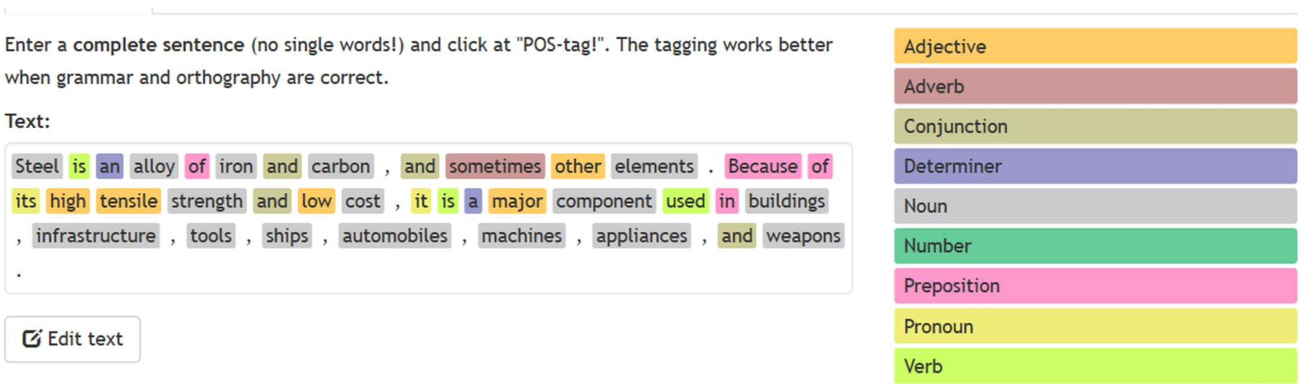
1. Exercise 1 (Lexical Analysis)

Using Wikipedia as a source, let's start by looking into the lexical categories (POS Tags) behind the words of some sentences.

'Steel is an alloy of iron and carbon, and sometimes other elements. Because of its high tensile strength and low cost, it is a major component used in buildings, infrastructure, tools, ships, automobiles, machines, appliances, and weapons.'

1.1 POS tag the sentences above.

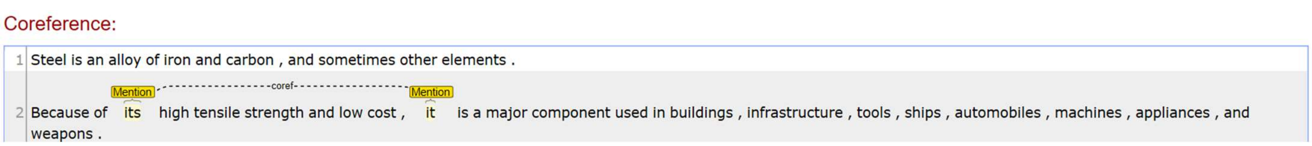
Steel is an alloy of iron and carbon , and sometimes other elements . Because of its high tensile
NN VBZ DT NN IN NN CC NN CC RB JJ NNS IN IN PRP$ JJ JJ
strength and low cost , it is a major component used in buildings, infrastructure, tools , ships , automobiles,
NN CC JJ NN PRP VBZ DT JJ NN VBN IN NNS NN NNS NNS NNS
machines, appliances, and weapons
NNS NNS CC NNS



1-1 POS tagging result on corenlp

1.2 Identify the pronominal co-references.

Co-reference: Because of its high … cost, it is …



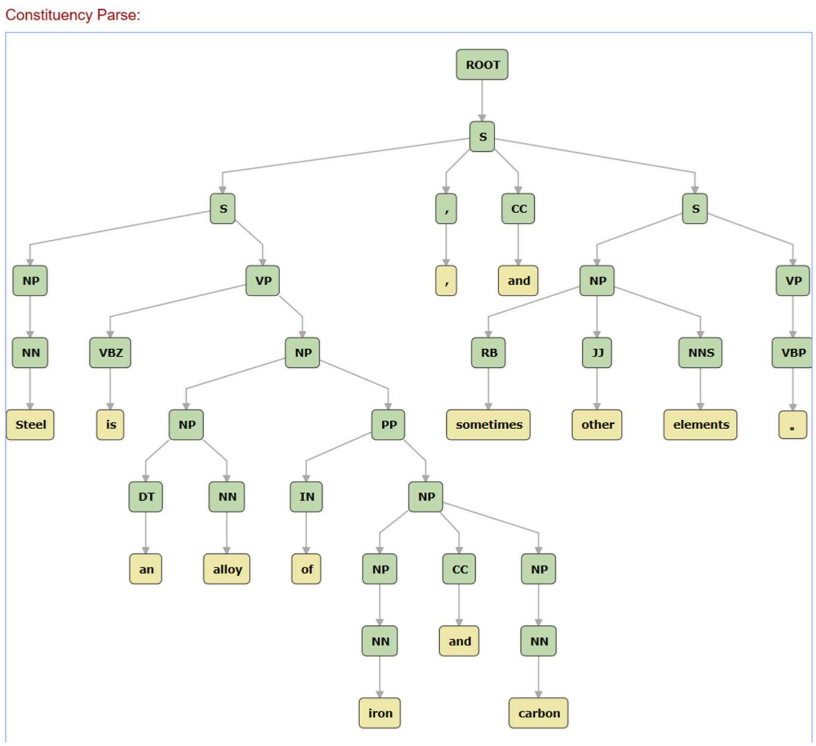
1-2 co-reference identification result on corenlp

These 2 references are both referencing 'Steel'.

2. Exercise 2 (C-Structures)

Using the Stanford Core NLP online:

2.1 Plot the constituency (phrase) structure of the sentence: ‘ Steel is an alloy of iron and carbon, and sometimes other elements’.



2-1 constituency structure on corenlp

2.2 In the class, we saw that the nominal phrasal nodes (NPs) correspond to ‘molecules of meaning’. Please list them for that sentence.

We just need to find all nodes labeled with ‘NP’ in pic 2-1. Nominal phrasal nodes are: Steel, an alloy, iron, carbon, iron and carbon, an alloy of iron and carbon, sometimes other elements

2.3 List the coordinations within that sentence.

There are 2 coordinations in the sentence, colored in red in the copy below:
Steel is an alloy of iron and carbon, and sometimes other elements

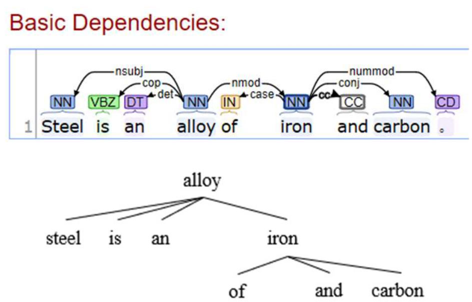
3. Exercise 3 (Dependencies - Exploring new territories)

In the class we did not talk about dependency parsing and dependency structures. They are a complementary perspective to constituency parsing.

3.1 Learn about dependency parsing (plenty of resources online). What is the difference between these two types of representation? What is emphasised by each representation?

A constituency parse tree breaks a text into sub-phrases. Non-terminals in the tree are types of phrases, the terminals are the words in the sentence, and the edges are unlabeled.
A dependency parse connects words according to their relationships. Each vertex in the tree represents a word, child nodes are words that are dependent on the parent, and edges are labeled by the relationship.

3.2 Draw the dependency structure of the following sentence: ‘Steel is an alloy of iron and carbon’.

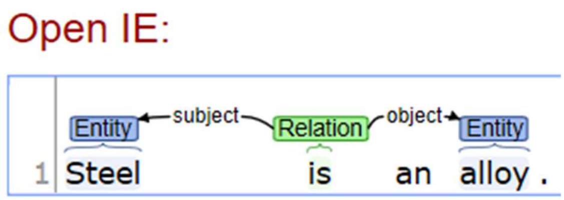


3-1 dependency structure of the sentence

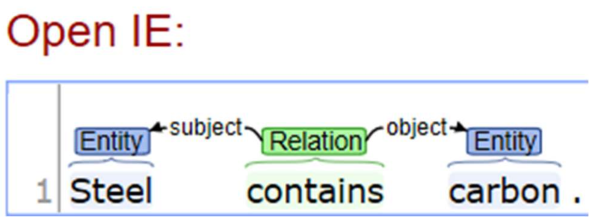
4. Exercise 4 (Open IE - Semantics)

Dependency structures are a form of syntactic representation which is almost at a semantic level of representation (meaning that they are almost at a predicate argument structure level).

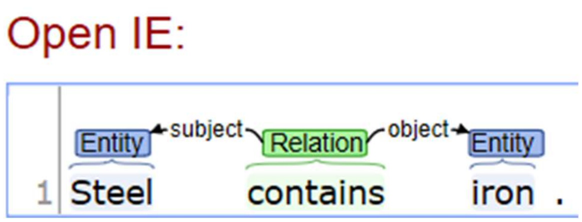
4.1 Now, let's shift gears to the semantic level and use Open Information Extraction (OpenIE) to extract the predicate-argument structure (s p o) of the following sentences (use CoreNLP for this): 'Steel is an alloy.' 'Steel contains carbon.' 'Steel contains iron.'



4-1 predicate-argument structure of 'Steel is an alloy.'



4-2 predicate-argument structure of 'Steel contains carbon.'



4-3 predicate-argument structure of 'Steel contains iron.'

4.2 Represent the triples above using Prolog.

is (steel, alloy).
contain(steel, carbon).
contain(steel, iron).

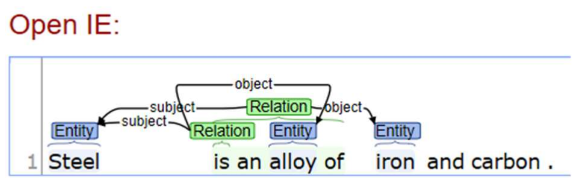
4.3 Represent the triples above using RDF.

< steel><is an><alloy>
< steel><contains><carbon>
< steel><contains><iron>

4.4 Formalize the axioms using Description Logics.

Steel ⊆ is.alloy
Steel ⊆ contains.carbon
Steel ⊆contains.iron

4.5 Now, analyse what happens when you get a slightly more complicated sentence into the system: 'Steel is an alloy of iron and carbon.'



4-4 predicate-argument structure of ‘Steel is an alloy of iron and carbon.’

In this step We can see that ‘and carbon’ is not covered as sentence becomes more complicated.

Prolog:

Since the sentence is not simple enough to be written in a single prolog fact, we can only convert it to 3 simpler sentences and write them in prolog as that in 4.2.

RDF:

RDF shall also be the same as that in 4.3 because format and content of RDF is rather dependent on content of prolog.

DL:

Since we can only extract one predicate from this complex sentence, yet in that case object is not atomized to nouns, DL formula cannot be written directly. The only way is also converting it to simpler sentences and write it like that in 4.4.

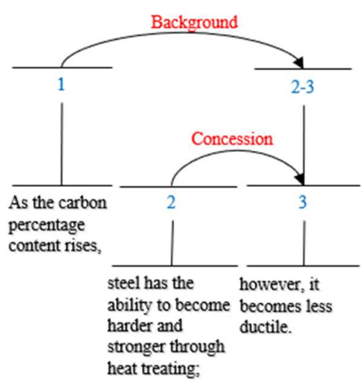
5. Exercise 5 (Complex Open IE, Rhetorical Structures)

Let's push OpenIE to the next level. Please take a look on how a different system, Graphene, does OpenIE, by disembedding phrases and clauses (i.e. by doing sentence splitting) and assigning rhetorical relations to it. In case you need more detail, please refer to 5, 6. Now, for the following sentence:
'As the carbon percentage content rises, steel has the ability to become harder and stronger through heat treating; however, it becomes less ductile.'

5.1 Identify the nucleus and the satellites.

Nucleus: steel has the ability to become harder and stronger through heat treating; however, it becomes less ductile.
Satellite: As the carbon percentage content rises,

5.2 Draw the diagram with the rhetorical relations.



5-1 rhetorical relation graph of target sentence

5.3 Which relations are hypotactic or paratactic?

Background and concession are both paratactic.

5.4 Write the output for that sentence using the RDF-NL notation of Graphene. It is not necessary to run Graphene (but you can compare your representation with its output if you want).

Name: output_re_1 →
#As the carbon percentage content rises , steel has the ability to become harder and stronger through heat treating ; however , it becomes less ductile .
ed9b4ce97c6942e7bc30e5cae92c62af 0 Steel has the ability to become harder and stronger through heat
    L:IDENTIFYING_DEFINITION c9981fa624114dc597654fda140cc510
    L:BACKGROUND fac56e0386654502876eedf9041bc703
    L:CONTRAST 057bac6e67974325961f665894e77ad9
e5bc525a2ced423d9e03050d47271473 1 Heat is treating
bfd40a81d12c4cc281f51b1d1bb0549e 1 The carbon percentage content rises
508da0e65ef643b78b508735fa0ef94c 0 It becomes less ductile
    L:CONTRAST 525eec5a203146c1990a448d8a40bed0

6. Exercise 6 (Taxonomies, Thesauri)

Now let's go even more semantic! Let's see how WordNet can help us to bring more structured domain knowledge. Let's take a technical term, related to the crystallographic structure/phase transition of steel: martensite and austenite. Using WordNet online:

6.1 List the WordNet glosses for these two words.

gloss for martensite:

a solid solution of carbon in alpha-iron that is formed when steel is cooled so rapidly that the change from austenite to pearlite is suppressed; responsible for the hardness of quenched steel

gloss for austenite:

a solid solution of ferric carbide or carbon in iron; cools to form pearlite or martensite

6.2 List the Taxonomic/Hypernym chain up to the top for these two terms.

Hypernym for martensite:

{15076272} <noun.substance>[27] S: (n) solid_solution#1 (solid_solution%1:27:00::), primary_solid_solution#1 (primary_solid_solution%1:27:00::) (a homogeneous solid that can exist over a range of component chemicals; a constituent of alloys that is formed when atoms of an element are incorporated into the crystals of a metal)

Hypernym for austenite:

{15076272} <noun.substance>[27] S: (n) solid_solution#1 (solid_solution%1:27:00::), primary_solid_solution#1 (primary_solid_solution%1:27:00::) (a homogeneous solid that can exist over a range of component chemicals; a constituent of alloys that is formed when atoms of an element are incorporated into the crystals of a metal)

6.3 List their sibling terms.

Sibling terms for martensite:

{15076272} <noun.substance>[27] S: (n) solid_solution#1 (solid_solution%1:27:00::), primary_solid_solution#1 (primary_solid_solution%1:27:00::) (a homogeneous solid that can exist over a range of component chemicals; a constituent of alloys that is formed when atoms of an element are incorporated into the crystals of a metal)

1. {14793921} <noun.substance>[27] S: (n) austenite#1 (austenite%1:27:00::) (a solid solution of ferric carbide or carbon in iron; cools to form pearlite or martensite)
2. {14883352} <noun.substance>[27] S: (n) ferrite#1 (ferrite%1:27:00::) (a solid solution in which alpha iron is the solvent)
3. {14972515} <noun.substance>[27] S: (n) **martensite#1 (martensite%1:27:00::)** (a solid solution of carbon in alpha-iron that is formed when steel is cooled so rapidly that the change from austenite to pearlite is suppressed; responsible for the hardness of quenched steel)
4. {15036554} <noun.substance>[27] S: (n) double salt#1 (double_salt%1:27:00::) (a solution of two simple salts that forms a single substance on crystallization)

Sister terms for austenite:

{15076272} <noun.substance>[27] S: (n) solid_solution#1 (solid_solution%1:27:00::), primary_solid_solution#1 (primary_solid_solution%1:27:00::) (a homogeneous solid that can exist over a range of component chemicals; a constituent of alloys that is formed when atoms of an element are incorporated into the crystals of a metal)

1. {14793921} <noun.substance>[27] S: (n) **austenite#1 (austenite%1:27:00::)** (a solid solution of ferric carbide or carbon in iron; cools to form pearlite or martensite)
2. {14883352} <noun.substance>[27] S: (n) ferrite#1 (ferrite%1:27:00::) (a solid solution in which alpha iron is the solvent)
3. {14972515} <noun.substance>[27] S: (n) martensite#1 (martensite%1:27:00::) (a solid solution of carbon in alpha-iron that is formed when steel is cooled so rapidly that the change from austenite to pearlite is suppressed; responsible for the hardness of quenched steel)

4. {15036554} <noun.substance>[27] S: (n) double salt#1 (double_salt%1:27:00::) (a solution of two simple salts that forms a single substance on crystallization)

6.4 For the word temper, how many synsets do we have? Which senses are related to steel? What are its synonyms? Would you consider these perfect or near-synonyms?

There are 4 sysnets for noun senses and 5 for verb senses, altogether 9 sysnets;

{05028701} <noun.attribute>[07] S: (n) **temper#4 (temper%1:07:01::)**, toughness#3 (toughness%1:07:03::) (the elasticity and hardness of a metal object; its ability to absorb considerable energy before cracking)

{00303837} <verb.change>[30] S: (v) anneal#1 (anneal%2:30:00::), **temper#1 (temper%2:30:04::)**, normalize#3 (normalize%2:30:03::) (bring to a desired consistency, texture, or hardness by a process of gradually heating and cooling) *"temper glass"*

{00303656} <verb.change>[30] S: (v) **temper#2 (temper%2:30:02::)**, harden#3 (harden%2:30:04::) (harden by reheating and cooling in oil) *"temper steel"*

These senses are related to steel.

Synonyms:

{05027524} <noun.attribute>[07] S: (n) elasticity#1 (elasticity%1:07:00::), snap#8 (snap%1:07:00::) (the tendency of a body to return to its original shape after it has been stretched or compressed) "the waistband had lost its snap"

1. {05027863} <noun.attribute>[07] S: (n) resilience#1 (resilience%1:07:00::), resiliency#2 (resiliency%1:07:00::) (the physical property of a material that can return to its original shape or position after deformation that does not exceed its elastic limit)
2. {05028147} <noun.attribute>[07] S: (n) bounce#1 (bounce%1:07:00::), bounciness#1 (bounciness%1:07:00::) (the quality of a substance that is able to rebound)
3. {05028317} <noun.attribute>[07] S: (n) give#1 (give%1:07:00::), spring#5 (spring%1:07:00::), springiness#1 (springiness%1:07:00::) (the elasticity of something that can be stretched and returns to its original length)
4. {05028511} <noun.attribute>[07] S: (n) stretch#7 (stretch%1:07:00::), stretchiness#1 (stretchiness%1:07:00::), stretchability#1 (stretchability%1:07:00::) (the capacity for being stretched)
5. {05028906} <noun.attribute>[07] S: (n) elasticity of shear#1 (elasticity_of_shear%1:07:00::) (the elasticity of a body that has been pulled out of shape by a shearing force)

These are near-synonyms, because these senses do not cover 'hardness' in sense of 'temper'.

7.Exercise 7 (Frame Semantics - Exploring Further)

Now let's take a look into the semantic representation of verbs using Frame Semantics. Using VerbNet as a starting point, for the verbs melt and oxidize:

7.1 Describe the frame semantics for these verbs as listed by VerbNet, PropBank and FrameNet.

For 'melt':

VerbNet:

Agent [+animate | +machine]

Product

Material

Propbank:

Arg0-PAG: agent (vnrole: 26.5-Agent, 45.4-Agent)

Arg1-PPT: thing melted (vnrole: 26.5-Material, 45.4-Patient)

FrameNet:

Cause_change_of_phase(A Cause or Agent causes a Patient to undergo a change of phase. The Result of the change may be given, along with the Initial_state and the Circumstances under which the change can occur. Note that this frame contrasts with Cause_change_of_consistency in that this frame describes causation of a change of a Patient between different phases (i.e. solid to liquid or frozen to "unfrozen").)

Change_of_phase(In this frame a Patient undergoes a change of phase. Note that this frame contrasts with Change_of_consistency in that this frame describes a change of a Patient between different phases (i.e. solid to liquid or frozen to "unfrozen").)

For 'oxidize':

VerbNet:

Patient [+concrete]

Propbank:

Arg0-PAG: Cause of oxidizing, Scientist causing conversion

Arg1-PPT: Entity/element oxidizing

Arg2-PRD: End state, oxidized to what

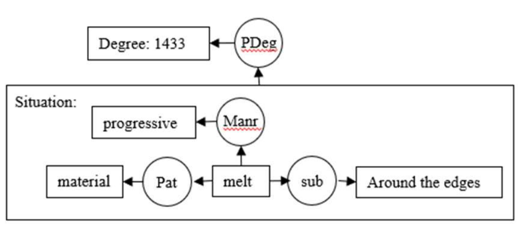
FrameNet:

Corroding_caused (In this frame, the Patient, an inorganic item, undergoes a chemical process caused by another entity, which renders it less useful, desirable or stable. The process may or may not be under the control of an Agent.)

7.2 How these representations compare?

Propbank will give fewer roles because they are more generalized, frames in Propbank are defined as prototypes; representations in FrameNet are larger in amount and more specific. For Verbnet, it only give schema for verbs, if we are looking a noun under this frame, we are likely to find only its verbal frame or even nothing in it.

7.3 Using the FrameNet semantics, draw the Conceptual Graph for the following sentence: 'At 1433 degrees, the material started to progressively melt around the edges.'



7-1 conceptual graph for target sentence

8. Exercise 8 (Ontologies - Description Logics)

Now let's use description logics to structure our domain into an ontology. You can decide if you prefer to do it using purely DL formulae or to use a supporting ontology editor 10 (recommended). You can learn to use protege using this tutorial 11.

Convert the following natural language statements into an ontology:

There are many types of heat treating processes available to steel. The most common are annealing, quenching, and tempering.

- Low-carbon steel: 0.05 to 0.30% carbon content.
- Medium-carbon steel: Approximately 0.3{0.6% carbon content. Balances ductility and strength and has good wear resistance; used for large parts, forging and automotive components.
- High-carbon steel: Approximately 0.60 to 1.00% carbon content. Very strong, used for springs, edged tools, and high-strength wires.
- Ultra-high-carbon steel: Approximately 1.25{2.0% carbon content. Steels that can be tempered to great hardness. Used for special purposes like knives, axles or punches. Most steels with more than 2.5% carbon content are made using powder metallurgy.

DL formulae:

HeatTreatingProcess ≡ annealing ∏ quenching ∏ tempering ∏ others

LowcarbonSteel ⊆ Steel

LowcarbonSteel ≡ ⩾5hasCarbonBaiscPointShare ∏ ⩽29 hasCarbonBaiscPointShare

MediumCarbonSteel ⊆ Steel

MediumCarbonSteel ≡ ⩾30hasCarbonBaiscPointShare ∏ ⩽59hasCarbonBaiscPointShare ∏ hasProperty.Ductility ∏ hasProperty.Strength ∏ hasProperty.WearResistance ∏ usedFor.LargePart ∏ usedFor.Forging ∏ usedForAutomotiveComponent

HighCarbonSteel ⊆ Steel

HighCarbonSteel ≡ ⩾60hasCarbonBaiscPointShare ∏ ⩽99hasCarbonBaiscPointShare ∏ hasProperty.Strength ∏ usedFor.Spring ∏ usedFor.EdgedTool ∏ usedFor.HighStrengthWire

UltraHighCarbonSteel ⊆ Steel

UltraHighCarbonSteel ≡ ⩾125hasCarbonBaiscPointShare ∏ ⩽199hasCarbonBaiscPointShare ∏ hasProperty. Hardness ∏ usedFor.SpecialPurpose

Knives ⊆ SpecialPurpose

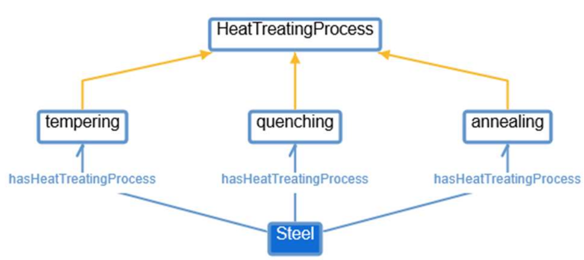
Axles ⊆ SpecialPurpose

Punches ⊆ SpecialPurpose

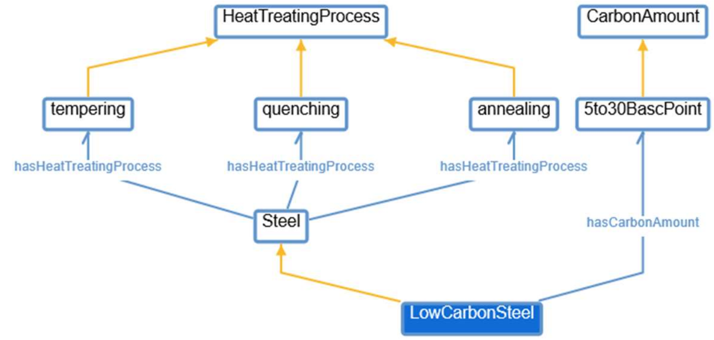
SuperHighCarbonSteel ⊆ Steel

SpecificSteel ≡ madeBy.PowderMetllurgy

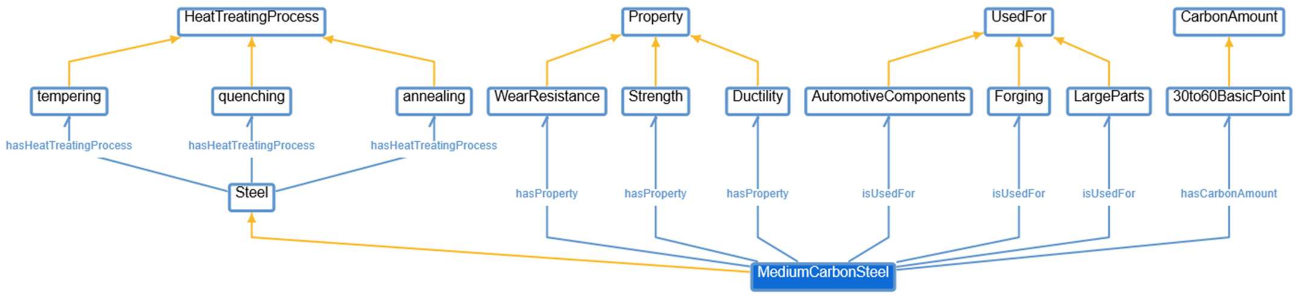
SuperHighCarbonSteel ≡ ⩾250hasCarbonBaiscPointShare



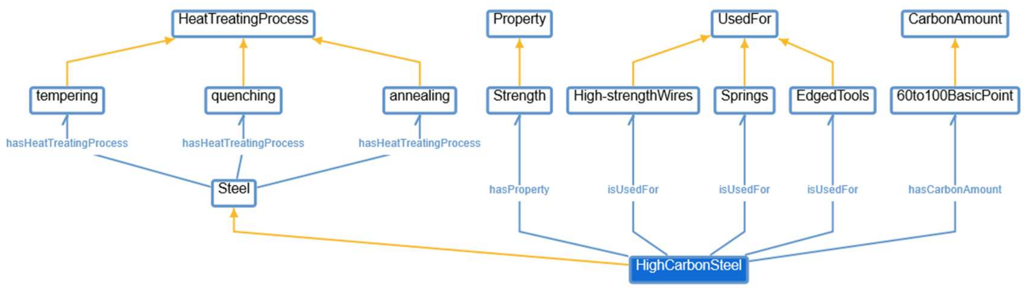
8-1 OWL graph for Steel



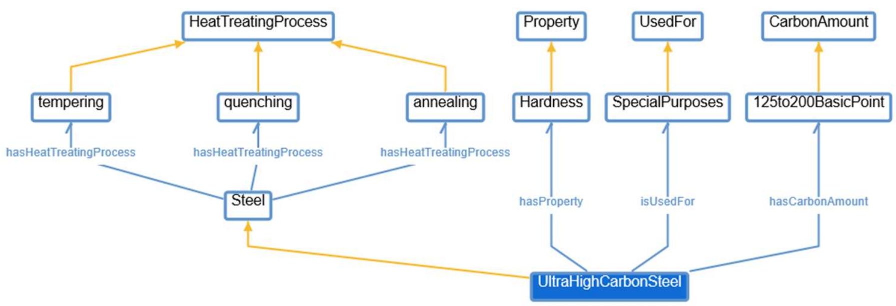
8-2 OWL graph for Low Carbon Steel



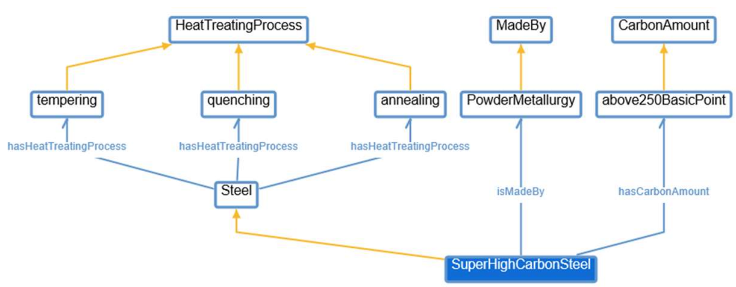
8-3 OWL graph for Medium Carbon Steel



8-4 OWL graph for High Carbon Steel



8-5 OWL graph for Ultra High Carbon Steel



8-6 OWL graph for Super High Carbon Steel

9. Exercise 9 (Inductive & Deductive Reasoning)

9.1 The first task is using ILP to learn the definitions of each type of crystallographic structure given the sets of positive and negative examples. You should output rules for the following structures, where X is a variable corresponding to the temperature and Y a variable corresponding to the carbon concentration.

AUSTENITE(PARAM_0, PARAM_1) :- MEDIUM_CARBON(PARAM_1), MEDIUM_TEMPERATURE(PARAM_0).
CEMENTITE(PARAM_0, PARAM_1) :- HIGH_CARBON(PARAM_1), MEDIUM_TEMPERATURE(PARAM_0).
FERRITE(PARAM_0, PARAM_1) :- LOW_CARBON(PARAM_1), MEDIUM_TEMPERATURE(PARAM_0).
PEARLITE(PARAM_0, PARAM_1) :- MEDIUM_CARBON(PARAM_1), LOW_TEMPERATURE(PARAM_0).

9.2 The following background knowledge describes whether a generic configuration X is liquid or solid according to the temperature. The second task requires you to perform deductive reasoning by using the learned predicates and the additional background knowledge together. With the inductive + deductive process you should answer the following query: Are the materials solid or liquid?

```
?- ferrite(X,Y),solid(X).
X = 1000,
Y = 0.4 ;
X = 900,
Y = 0.4 ;
false.

?- ferrite(X,Y),liquid(X).
false.
```

```
?- austenite(X,Y),solid(X).
X = 1000,
Y = 0.78 ;
X = 900,
Y = 0.78 ;
false.

?- austenite(X,Y),liquid(X)
false.
```

```
?- pearlite(X,Y),solid(X).
X = 700,
Y = 0.78.

?- pearlite(X,Y),liquid(X).
false.
```

```
?- cementite(X,Y),solid(X).
X = 1000,
Y = 2.0 ;
X = 900,
Y = 2.0 ;
false.

?- cementite(X,Y),liquid(X)
false.
```

9-1~9-4: prolog result for querying states of 4 generic configuration X.

With prolog deduction results shown in pic 9-1 to 9-4, we can see that all 4 generic configurations are in solid state.

9.3 Depending on the type of engineering application, different material properties are required. The third task requires you to infer the material properties of the different types of material. By discretizing Figure 2, we can derive additional background knowledge. Given the additional background knowledge and the learned definitions you are required to answer the queries below. Also please provide a supporting explanation. (this again has to be done by using deduction).

```
?- ferrite(X,Y), high_hardness(Y).
false.
```

9-5 query result for 'Does Ferrite have high hardness? Why?'

We can see that ferrite does not have high hardness, which we can directly see from pic 9-5..

```
?- (ferrite(X,Y), high_ductility(Y), cementite(Z,W), medium_ductility(W));
|     (ferrite(X,Y), high_ductility(Y), cementite(Z,W), low_ductility(W));
|     (ferrite(X,Y), medium_ductility(Y), cementite(Z,W), low_ductility(W)).
X = Z, Z = 1000,
Y = 0.4,
W = 2.0 ;
X = 1000,
Y = 0.4,
Z = 900,
W = 2.0 ;
X = 900,
Y = 0.4,
Z = 1000,
W = 2.0 ;
X = Z, Z = 900,
Y = 0.4,
W = 2.0 ;
false.
```

9-6 query result for 'Is Ferrite more ductile than Cementite? Why?'

We can see that Ferrite is always more ductile than Cementite, since as long as the condition meets the requirement that Ferrite and Cementite exists simultaneously, Ferrite will always contain less carbon than Cementite, thus being more ductile.

```
?- ferrite(X,Y),low_tensile_strength(Y).
X = 1000,
Y = 0.4 .

?- austenite(X,Y),low_tensile_strength(Y).
false.

?- cementite(X,Y), low_tensile_strength(Y)
false.

?- pearlite(X,Y), low_tensile_strength(Y).
false.
```

9-7 query result for 'Which material has low tensile strength? Why?'

Since only query ferrite(X,Y), low_tensile_strength (Y) is returning results and all other queries is getting us false, we know that only ferrite has low tensile strength.

9.4 KB for 9.2 and 9.3

```
high_temperature(1535).
medium_temperature(1000).
medium_temperature(900).
low_temperature(700).
high_carbon(2.0).
medium_carbon(0.78).
low_carbon(0.4).

austenite(Param_0, Param_1) :- medium_carbon(Param_1), medium_temperature(Param_0).
cementite(Param_0, Param_1) :- high_carbon(Param_1), medium_temperature(Param_0).
ferrite(Param_0, Param_1) :- low_carbon(Param_1), medium_temperature(Param_0).
pearlite(Param_0, Param_1) :- medium_carbon(Param_1), low_temperature(Param_0).

liquid(X) :- high_temperature(X).
solid(X) :- medium_temperature(X).
solid(X) :- low_temperature(X).

high_hardness(Y) :- high_carbon(Y).
medium_hardness(Y) :- medium_carbon(Y).
low_hardness(Y) :- low_carbon(Y).
high_ductility(Y) :- low_carbon(Y).
medium_ductility(Y) :- medium_carbon(Y).
low_ductility(Y) :- high_carbon(Y).
high_tensile_strength(Y) :- medium_carbon(Y).
medium_tensile_strength(Y) :- high_carbon(Y).
low_tensile_strength(Y) :- low_carbon(Y).
```