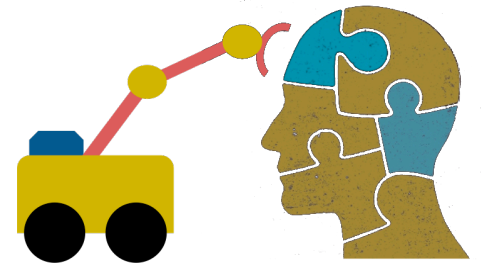


9.3.2



# First-Order Logic (FOL) part 2

# Overview

- We'll first give some examples of how to translate between FOL and English
- Then look at modelling family relations in FOL
- And finally touch on a few other topics

# Translating English to FOL

**Every gardener likes the sun**

$$\forall x \text{ gardener}(x) \rightarrow \text{likes}(x, \text{Sun})$$

**All purple mushrooms are poisonous**

$$\forall x (\text{mushroom}(x) \wedge \text{purple}(x)) \rightarrow \text{poisonous}(x)$$

**No purple mushroom is poisonous (two ways)**

$$\neg \exists x \text{ purple}(x) \wedge \text{mushroom}(x) \wedge \text{poisonous}(x)$$

$$\forall x (\text{mushroom}(x) \wedge \text{purple}(x)) \rightarrow \neg \text{poisonous}(x)$$

# English to FOL: Counting



Use = predicate to identify different individuals

- There are at least two purple mushrooms

$$\exists x \exists y \text{ mushroom}(x) \wedge \text{purple}(x) \wedge \text{mushroom}(y) \wedge \text{purple}(y) \wedge \neg(x=y)$$

- There are exactly two purple mushrooms

$$\begin{aligned} &\exists x \exists y \text{ mushroom}(x) \wedge \text{purple}(x) \wedge \text{mushroom}(y) \wedge \text{purple}(y) \wedge \neg(x=y) \wedge \\ &\forall z (\text{mushroom}(z) \wedge \text{purple}(z)) \rightarrow ((x=z) \vee (y=z)) \end{aligned}$$

Saying there are 802 different Pokemon is hard!

Direct use of FOL is not for everything!

# Translating English to FOL



What do these mean?

- You can fool *some of* the people *all of* the time
- You can fool *all of* the people *some of* the time

# Translating English to FOL



## What do these mean?

Both English statements are ambiguous

- **You can fool *some of the people* *all of the time***

There is a nonempty subset of people so easily fooled that you can fool that subset every time\*

For any given time, there is a non-empty subset at that time that you can fool

- **You can fool *all of the people* *some of the time***

There are one or more times when it's possible to fool everyone\*

Each individual can be fooled at some point in time

\* Most common interpretation, I think

# Some terms we will need



- **person(x)**: True iff x is a person
- **time(t)**: True iff t is a point in time
- **canFool(x, t)**: True iff x can be fooled at time t

Note: *iff* = *if and only if* =  $\leftrightarrow$

# Translating English to FOL



**You can fool *some of* the people *all of* the time**

There is a nonempty group of people so easily fooled that you can fool that group every time\*

≡ There's (at least) one person you can fool every time

$$\exists x \forall t \text{ person}(x) \wedge \text{time}(t) \rightarrow \text{canFool}(x, t)$$

For any given time, there is a non-empty group at that time that you can fool

≡ For every time, there's a person at that time that you can fool

$$\forall t \exists x \text{ person}(x) \wedge \text{time}(t) \rightarrow \text{canFool}(x, t)$$

\* Most common interpretation, I think



# Translating English to FOL



**You can fool *all of* the people *some of* the time**

There's at least one time when you can fool everyone\*

$$\exists t \forall x \text{time}(t) \wedge \text{person}(x) \rightarrow \text{canFool}(x, t)$$

Everybody can be fooled at some point in time

$$\forall x \exists t \text{person}(x) \wedge \text{time}(t) \rightarrow \text{canFool}(x, t)$$

\* Most common interpretation, I think

# Representation Design



- Many options for representing even a simple fact, e.g., something's color as red, green or blue, e.g.:
  - green(kermit)
  - color(kermit, green)
  - hasProperty(kermit, color, green)
- Choice can influence how easy it is to use
- Last option of representing properties & relations as triples used by modern knowledge graphs
  - Easy to ask: What color is Kermit? What are Kermit's properties?, What green things are there? Tell me everything you know, ...

# Simple genealogy KB in FOL



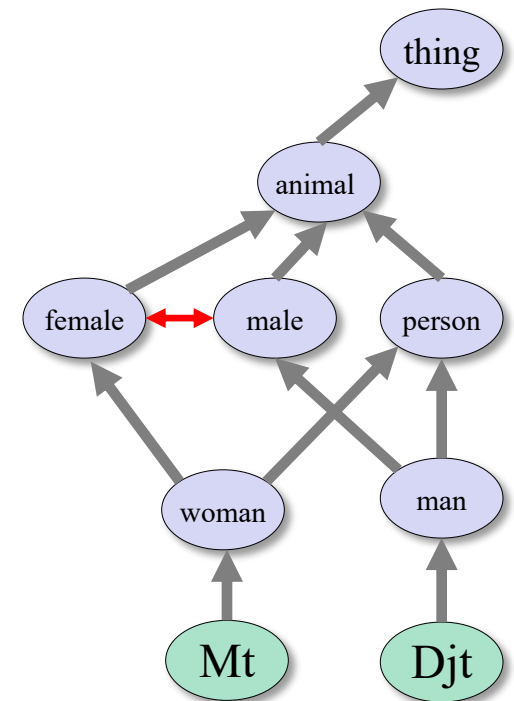
**Design a knowledge base using FOL that**

- Has facts of immediate family relations, e.g., spouses, parents, etc.
- Defines more complex relations (ancestors, relatives)
- Detect conflicts, e.g., you are your own parent
- Infers relations, e.g., grandparent from parent
- Answers queries about relationships between people

# How do we approach this?



- Design an initial ontology of types, e.g.
  - person, animal, man, woman, ...
- Types form a taxonomy or lattice\*, e.g.
  - $\text{person}(X) \Leftrightarrow \text{man}(X) \vee \text{woman}(Y)$
  - $\text{man}(X) \Leftrightarrow \text{person}(X) \wedge \text{male}(X)$
  - $\text{woman}(X) \Leftrightarrow \text{person}(X) \wedge \text{female}(X)$
  - $\text{female}(X) \Leftrightarrow \sim \text{male}(X)$
- Make assertions about individuals, e.g.
  - $\text{man}(\text{Djt})$
  - $\text{woman}(\text{Mt})$



\* In a [lattice](#), objects can have multiple immediate types

# Extend with relations and constraints

- Simple two argument relations, e.g.
  - spouse, has\_child, has\_parent
- Add general constraints to relations, e.g.
  - $\text{spouse}(X,Y) \Rightarrow \sim (X = Y)$
  - $\text{spouse}(X,Y) \Rightarrow \text{person}(X) \wedge \text{person}(Y)$
  - $\text{spouse}(X,Y) \Rightarrow (\text{man}(X) \wedge \text{woman}(Y)) \vee (\text{woman}(X) \wedge \text{man}(Y))^*$
- Add FOL sentences for inference, e.g.
  - $\text{spouse}(X,Y) \Leftrightarrow \text{spouse}(Y,X)$
- Add instance data
  - e.g.,  $\text{spouse}(\text{Djt}, \text{Mt})$

\* Note this constraint is a traditional one than no longer holds

# Example: A simple genealogy KB in FOL

## Predicates:

- parent(x, y), child(x, y), father(x, y), daughter(x, y), etc.
- spouse(x, y), husband(x, y), wife(x, y)
- ancestor(x, y), descendant(x, y)
- male(x), female(y)
- relative(x, y)

## Facts:

- husband(Joe, Mary), son(Fred, Joe)
- spouse(John, Nancy), male(John), son(Mark, Nancy)
- father(Jack, Nancy), daughter(Linda, Jack)
- daughter(Liz, Linda)
- etc.

# Example Axioms



$(\forall x,y) \text{parent}(x, y) \leftrightarrow \text{child}(y, x)$

$(\forall x,y) \text{father}(x, y) \leftrightarrow \text{parent}(x, y) \wedge \text{male}(x)$

$(\forall x,y) \text{mother}(x, y) \leftrightarrow \text{parent}(x, y) \wedge \text{female}(x)$

$(\forall x,y) \text{daughter}(x, y) \leftrightarrow \text{child}(x, y) \wedge \text{female}(x)$

$(\forall x,y) \text{son}(x, y) \leftrightarrow \text{child}(x, y) \wedge \text{male}(x)$

$(\forall x,y) \text{husband}(x, y) \leftrightarrow \text{spouse}(x, y) \wedge \text{male}(x)$

$(\forall x,y) \text{spouse}(x, y) \leftrightarrow \text{spouse}(y, x)$

...

# Axioms, definitions and theorems

- **Axioms**: facts and rules that capture (important) facts & concepts in a domain; used to prove **theorems**
  - Mathematicians dislike unnecessary (dependent) axioms, i.e., ones that can be derived from others
  - Dependent axioms can make reasoning faster, however
  - Choosing a good set of axioms is a design problem
- A **definition** of a predicate is of the form “ $p(X) \leftrightarrow \dots$ ” and can be decomposed into two parts
  - **Necessary** description: “ $p(x) \rightarrow \dots$ ”
  - **Sufficient** description “ $p(x) \leftarrow \dots$ ”
  - Some concepts have definitions (e.g., triangle) and some don't (e.g., person)



# More on definitions

Example: define  $\text{father}(x, y)$  by  $\text{parent}(x, y)$  &  $\text{male}(x)$

- **$\text{parent}(x, y)$**  is a necessary (but not sufficient) description of  $\text{father}(x, y)$

$$\text{father}(x, y) \rightarrow \text{parent}(x, y)$$

- **$\text{parent}(x, y) \wedge \text{male}(x) \wedge \text{age}(x, 35)$**  is a sufficient (but not necessary) description of  $\text{father}(x, y)$ :

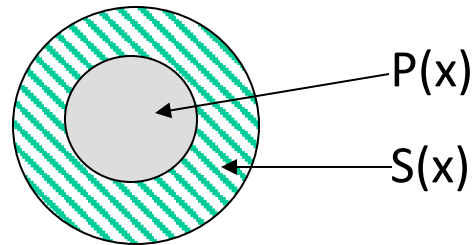
$$\text{father}(x, y) \leftarrow \text{parent}(x, y) \wedge \text{male}(x) \wedge \text{age}(x, 35)$$

- **$\text{parent}(x, y) \wedge \text{male}(x)$**  is a necessary and sufficient description of  $\text{father}(x, y)$

$$\text{parent}(x, y) \wedge \text{male}(x) \leftrightarrow \text{father}(x, y)$$

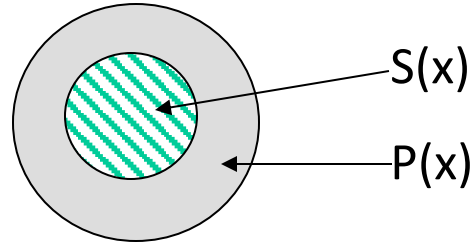
# Another way to look at necessary and sufficient

$S(x)$  is a necessary condition of  $P(x)$



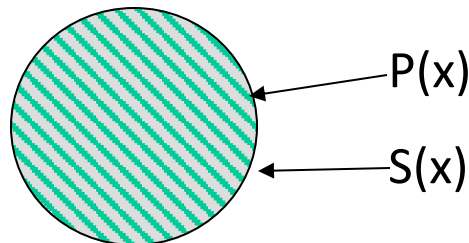
# all Ps are Ss  
 $(\forall x) P(x) \Rightarrow S(x)$

$S(x)$  is a sufficient condition of  $P(x)$



# all Ps are Ss  
 $(\forall x) P(x) \Leftarrow S(x)$

$S(x)$  is a necessary and sufficient condition of  $P(x)$



# all Ps are Ss  
# all Ss are Ps  
 $(\forall x) P(x) \Leftrightarrow S(x)$

# Higher-order logic

- FOL only lets us quantify over variables, and **variables can only range over objects**
- HOL allows us to quantify over relations, e.g.  
“two functions are equal iff they produce the same value for all arguments”

$$\forall f \forall g (f = g) \leftrightarrow (\forall x f(x) = g(x))$$

- E.g.: (quantify over predicates)

$$\forall r \text{transitive}( r ) \rightarrow (\forall xyz) r(x,y) \wedge r(y,z) \rightarrow r(x,z))$$

- More expressive, but reasoning is undecidable, in general



# Examples of FOL in use

- Semantics of W3C's [Semantic Web](#) stack (RDF, RDFS, OWL) is defined in FOL
- [OWL](#) Full is equivalent to FOL
- Other OWL profiles support a subset of FOL and are more efficient
- FOL oriented knowledge representation systems have many user friendly tools
- E.g.: Protégé for creating, editing and exploring OWL ontologies



schema.org

# Examples of FOL in use



Many practical approaches embrace the approach that “some data is better than none”

- The semantics of [schema.org](https://schema.org) is only defined in natural language text
- [Wikidata](https://wikidata.org)'s knowledge graph has a rich schema
  - Many constraint/logical violations are flagged with warnings
  - However, not all, see this Wikidata query that finds people who are their own mother or father

# Wikidata



“collaboratively edited structured dataset used by Wikimedia sister projects and others”

**Goal:** consolidate *knowable facts* for use in the 200+ Wikipedia sites and other Wikimedia resources (e.g., in Infoboxes)

**Integrates** Wikimedia sites & links entities to ~7K external identifier systems (e.g., [GeoNames ID](#))

**Used** by Google’s Knowledge Graph, digital assistants (Siri, Alexa), Wikipedia Infoboxes, etc.

English Finin Talk Preferences Beta Watchlist  
Contributions Log out

Main Page Discussion

Search Wikidata

The Wikimedia Foundation wants to make it easier for you to talk to other editors. We invite you to review our proposed direction.

open

## Welcome to Wikidata

the free knowledge base with 57,425,191 data items that anyone can edit.

[Introduction](#) • [Project Chat](#) • [Community](#)

### Welcome!

Wikidata is a free and open knowledge base that can be read and edited by both humans and machines.

Wikidata acts as central storage for the **structured data** of its Wikimedia sister projects including Wikipedia, Wikivoyage, Wiktionary, Wikisource, and others.

Wikidata also provides support to many other sites and services beyond just Wikimedia projects! The content of Wikidata is [available under a free license](#), exported using standard formats, and can be [interlinked to other open data sets](#) on the linked data web.

Print/export  
Create a book  
Download as PDF  
Printable version

Tools  
[What links here](#)

# Wikidata



**Knowledge graph** with ~1B statements about ~93M items  
Fine-grained **ontology** has ~2M types and ~5K properties  
Data exposed as **RDF** triples  
String values tagged with **language id**  
Query using a standard **SPARQL** query service  
Many community **tools** for search, visualization, update  
Semantic Web realized !

The screenshot shows the Wikidata Query Service interface. At the top, there's a 'Query Helper' section with filters: 'instance of' set to 'university', '+ Filter' set to 'located in the administrative territorial entity', and 'Baltimore'. Below this is a '+ Show' button and a 'Limit' section. The main area contains a SPARQL query:

```
1 # find universities in Baltimore and at their English names
2 Select ?U ?ULabel where
3 { ?U wdt:P31 wd:Q3918;
4   wdt:P131 wd:Q5092
5 SERVICE wikibase:label
6 {bd:serviceParam wikibase:language "en"}
7 }
8
```

Below the query, there's a play button and a refresh button. At the bottom, there's a status bar showing '16 results in 114 ms' and buttons for 'Code', 'Download', and 'Link'. The results table has two columns: 'U' and 'ULabel'.

U	ULabel
<a href="#">wd:Q16982520</a>	UMBC College of Arts, Humanities and Social Sciences
<a href="#">wd:Q16982524</a>	UMBC College of Engineering and Information Technology

# Wikidata Entity

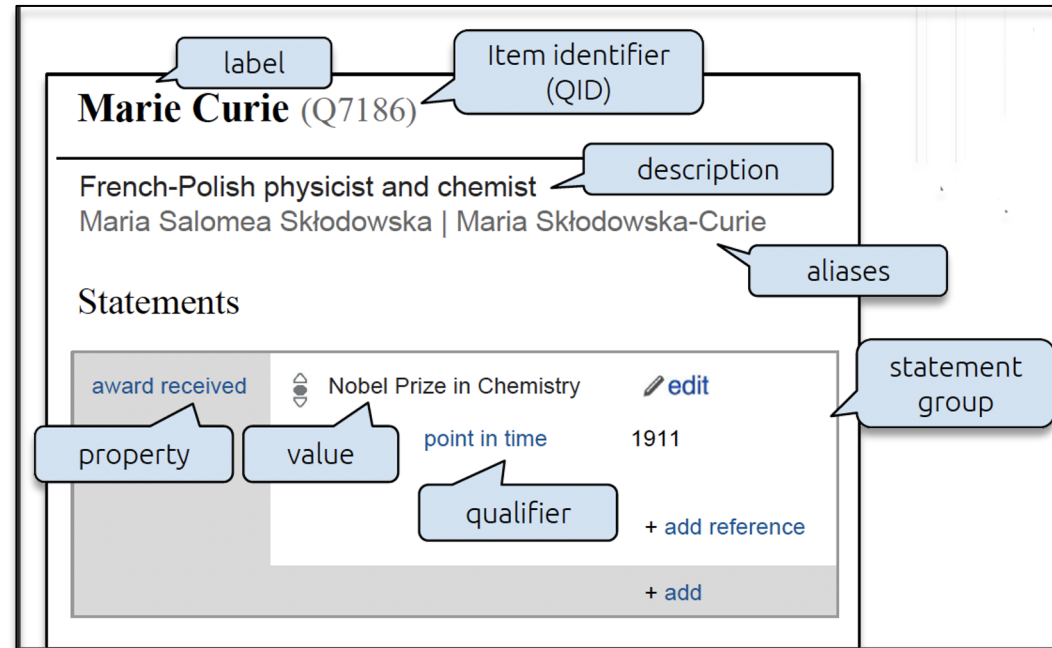
Each Wikipedia entity has

**Unique ID**, e.g. [Q7186](#)

A **label** (canonical name),  
short **description**,  $\geq 0$  **aliases** in a  
set of languages

One or more **types** (e.g., [Q5](#))

Collection of **statements** with  
optional qualifiers, references



Marie Curie is a human:

[wd:Q7186 wdt:P31 wd:Q5]



# Ontology exploration

How can we understand an ontology with more than two million types?

wdtaxonomy is a useful tool for exploring the ontology

Given a type (e.g, Q3918, university) you can quickly see

- Subtypes or supertypes (immediate or inferred),

- Number of instances (immediate or inferred),

- Direct instances

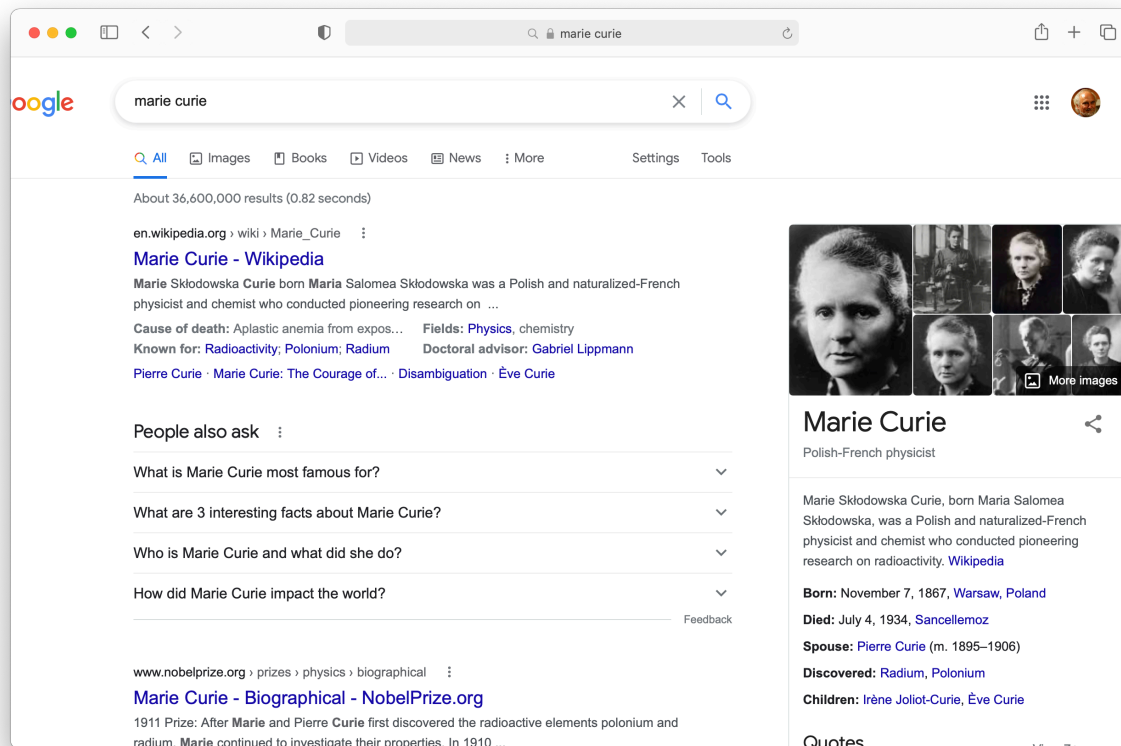
- Number Wikimedia sites it's in

Implemented in javascript with a command line script

```
$$ wdtaxonomy Q3918 -c -t
university (Q3918) •163 x15380 ↑
├─Universities in Germany (Q212462) •2
├─national university (Q265662) •11 x73
├─National University (Q366354) •5
├─Imperial universities of Japan (Q562092) •12
├─Byzantine university (Q622870) •4
├─college and university rankings (Q847843) •23 x45 ↑
├─public university (Q875538) •39 x974 ↑
├─private university (Q902104) •32 x846 ↑
├─new university (Q987075) •4 x1
├─Red brick university (Q1202123) •11
├─??? (Q1305046) •2
├─institute of technology (Q1371037) •20 x325
├─veterinary medicine school (Q1384955) •5 x28
├─online university (Q1407393) •4 x10 ↑
├─virtual university (Q1755248) •8 x11
├─online university (Q1407393) •4 x10 ↑ ...
├─comprehensive university (Q1767829) •2 x6
├─plate glass university (Q1902446) •8
├─medical university (Q1916585) •1 x9 ↑
├─??? (Q2073922) •1
├─pontifical university (Q2120466) •18 x37 ↑↑
├─Corporate university (Q2278672) •6
├─ancient university (Q2667285) •9 x1
├─central university (Q3351682) •12 x2
├─collegiate university (Q3354859) •9 x12
├─deemed university (Q3520135) •6 x16
├─university in France (Q3551775) •3 x75 ↑
├─Istituto superiore per le industrie artistiche (Q3803831) •2 x4
├─??? (Q3803846) •1 x2
├─Smolny Institute for Noble Maidens (Q4432880) •1
├─??? (Q4475845) •2
├─federal university (Q4481793) •3 x3
├─ecclesiastical university (Q5332280) •6 x2
├─labor universities (Q5690751) •1 x6
├─open university (Q6755402) •4 x1
├─Urban university (Q7900184) •2
├─??? (Q10387922) •1
├─international university (Q10829188) •3 x9
├─autonomous university (Q11057861) •2 x1
├─research university (Q15936437) •9 x224
├─Italian universities (Q20009854) •2
├─??? (Q20052016) •1 x2
├─Canciller de Universidad (Q21547263)
├─imperial university of the Russian Empire (Q28667313) •2 x12
├─universities in China (Q28700403) •1
├─Institute of National Importance (Q47531586) x1 ↑
├─campusuniversity (Q59537665) x3
├─Indiana University Bloomington Department of French and Italian (Q63441027)
├─Indiana University Department of French and Italian (Q63441251)
├─Indiana University Bloomington Department of History (Q63441447)
```

# Wikidata and Infoboxes

- Web search engines use custom **knowledge graphs** for infoboxes with information about queried items
- Wikidata is an open-source knowledge graph that shares roots with these
- They all draw on the same knowledge, like the ~300 Wikipedia & Wikimedia sites
- Such knowledge graphs are often used by language understanding systems, e.g., for entity linking (who is [Michael Jordan](#))



The screenshot shows a Google search for "marie curie". The search results include a link to the Wikipedia article for Marie Curie, a snippet of text describing her as a Polish and naturalized-French physicist and chemist, and a list of "People also ask" questions. On the right side, there is a detailed infobox for Marie Curie, including her name, title "Polish-French physicist", a grid of images, and biographical details such as her birth date (November 7, 1867, Warsaw, Poland), death date (July 4, 1934, Sancellemoz), spouse (Pierre Curie), and children (Irène Joliot-Curie, Ève Curie).

**Query Helper**

+ Filter

+ Show

Limit

father

mother

```
1 select ?Person ?PersonLabel where {  
2   ?Person wdt:P22|wdt:P25 ?Person .  
3   SERVICE wikibase:label {bd:serviceParam wikibase:language "en".}}
```

people who are their own mother or father

6 results in 2447 ms | </> Code | Download | Link

Person	PersonLabel
<a href="#">wd:Q99228816</a>	Augusta Sofia Dahlberg
<a href="#">wd:Q3489077</a>	Dave Lister
<a href="#">wd:Q13634217</a>	Marcus Asinius Marcellus
<a href="#">wd:Q15396074</a>	Önundur Oddsson
<a href="#">wd:Q64899126</a>	Carl Fredrik Oscar von Bahr
<a href="#">wd:Q98544781</a>	George Washington Crawford, I

people who are their own grandparent

# FOL Summary

- First order logic (FOL) introduces predicates, functions and quantifiers
- More expressive, but reasoning more complex
  - Reasoning in propositional logic is NP hard, FOL is semi-decidable
- Common AI knowledge representation language
  - Other KR languages (e.g., [OWL](#)) are often defined by mapping them to FOL
- FOL variables range over objects
  - HOL variables range over functions, predicates or sentences
- Some practical systems avoid enforcing rigid FOL constraints due to having noisy data

*Fín*