Annotation Schema

Classification:

1. Sluicing

Definition: sluicing is a type of ellipsis that occurs in both direct and indirect interrogative clauses. The ellipsis is introduced by a wh-expression, whereby in most cases, everything except the wh-expression is elided from the clause.

Example:

Yes! It is an example of sluicing ☺:

> **Phoebe ate** something, but she doesn't know what. (=what she ate)
> Where? (=where was Alan Turning born)

No! It is not an example of sluicing ☹:

> Turing studied where? (It is an example of inverted sentence. It has a fully grammatical structure without any ellipsis. Although it has a similar format as sluicing, we should not label it as sluicing.)

2. Anaphor

Definition: anaphora is the use of an expression whose interpretation depends upon another expression in context (its antecedent or postcedent). In our schema, anaphora is the use of an expression that depends specifically upon an antecedent expression.

If you can not fully interpret the meaning of an utterance solely based on itself without the prior context, we call it anaphor.

Example:

Yes! It is an example of anaphora ☺:

> It opened in November 1939. (What is it?)
> Which place does this airport serve? (Which airport is this?)
> And what is the date of publication? (The publication of what?)
> Who is another character of Harry Potter? (Which character has been mentioned?)

No! It is not an example of anaphora ☹:

> What's the population of tianjin? (We can answer the question without prior context.)
> 24 February 1955. (Although the date is describing the date of birth of Steve Jobs according to prior context, we could still interpret this data without of reviewing previous context. In other words, there is no ambiguity arising in this sentence.

3. Short Answer:

Definition: short answer (= answer fragments) is a type of ellipsis that occurs in answers to questions. In our schema, we define all answer utterances which is not a fully grammatical sentence to be short answer.

Example:

Yes! It is an example of short answer ☺:

> Q: Who walked the dog? A: Tom walked the dog. **- Subject noun as answer fragment**
> Q: Why will they resist our help? A: They will resist our help Due to excessive pride. **- Causal adjunct prepositional phrase as answer fragment**

No! It is not an example of short answer ☹:

> Q: Who walked the dog? A: Tom walked the dog.

4. Error:

Definition: error tag should only be annotated when there is a disagreement between the grounded knowledge graph and the utterances.

Example:

Yes! It is an example of short answer ☺:

> Q: How is the weather today? (However, the grounded relation is not talking about weather.)

No! It is not an example of short answer ☹:

> Q: When are Bill Gates bron? (Typo and grammatical error should not be annotated with error tags in our schema).

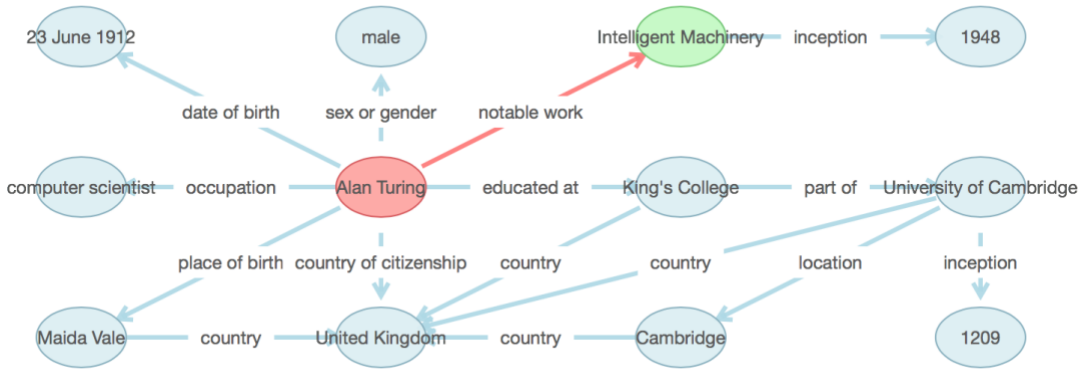
Labelling:

1. Additional information outside graph:

Definition: If parts of an utterance introduce additional information into the conversation which is not mentioned in the grounded knowledge graph, we call this part additional information outside the knowledge graph.

Example:

Given the knowledge graph and highlight:



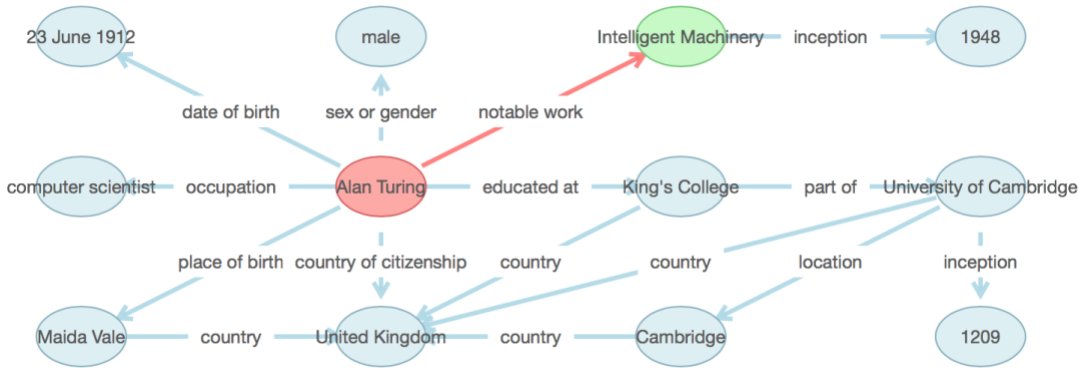
Yes! It is an example of Additional information outside graph ☺:

> Q: Can you tell me one of the notable work of Mr. Turning?
> A: Intelligent Machinery! Don't you head of the famous turning test?
> The underlining sentence is considered as addition information from outside of the knowledge graph, because there is no evidence that intelligent machinery is related to turning test.

2. Additional information within graph

Definition: If parts of an utterance introduce additional information into the conversation which is mentioned in the graph and such information is not highlighted, we call them additional information within the knowledge graph.

Given the knowledge graph and highlight:



Yes! It is an example of Additional information within graph ☺:

> Q: Can you tell me one of the notable work of Mr. Turning?
> A: Intelligent Machinery! Turing got this amazon idea in 1948.
> The underlining sentence is considered as addition information within of the knowledge graph, because the highlight relation doesn't contain year 1948. However, from the knowledge graph, we could know that Intelligent Machinery is incepted in 1948.

3. Conversation Control

Definition: Conversation Control label should be tagged to any pieces of utterance which functionally act as a connector to make the whole conversation more natural and coherent.

Yes! It is an example of Conversation Control ☺:

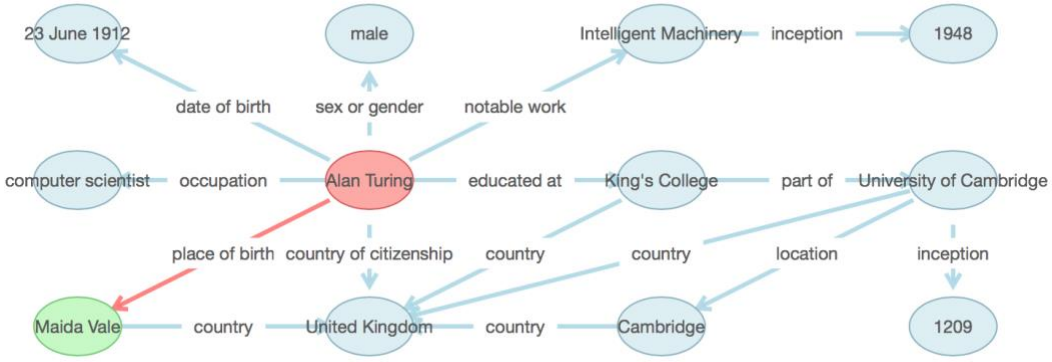
> Q: When was Alan Turing born?
> A: He was born in 1912.
> Q: And where?
> A: A town called Maida Vale.
> Q: Sorry, it is bit random. When was Alan's notable work, Intelligent Machinery incepted?
> A: 1948.

4. Totally Irrelevant

Definition: totally irrelevant tag should be used when pieces of utterance is totally irrelevant to the question answering conversation. By deleting such pieces of information, we will not influence the coherence of the conversation or the meaning of the utterance. In addition, this tag should only be considered

if the part of utterance is not labelled as other labels. In other words, labels, such as Conversation Control, will have higher precedence.

Example:

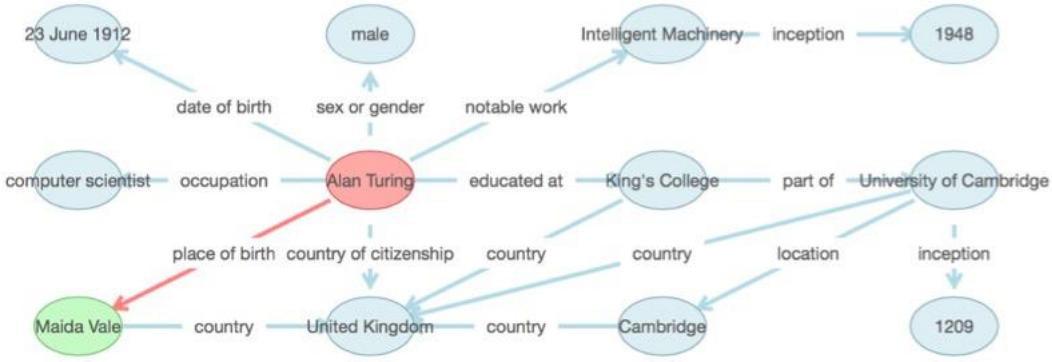


Yes! It is an example of Totally Irrelevant☺:

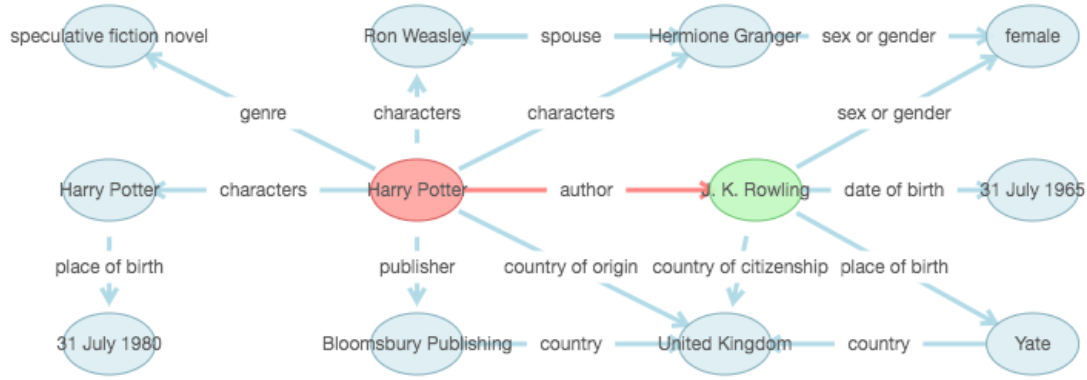
> 13,245,000. do you mind if we just do the next task and that's us finished or do you want to do another round?
>
> The underlining sentence is considered as totally irrelevant information since it is not directly related to our conversations.

Appendix A: All Grounded Graphs for this Tasks.

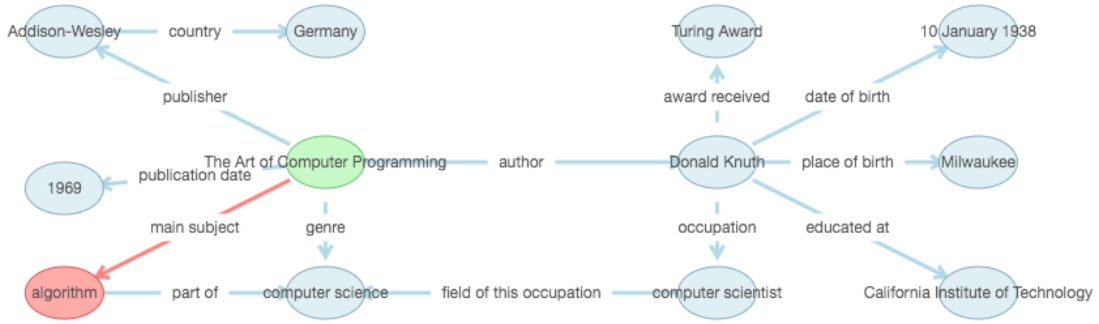
1.



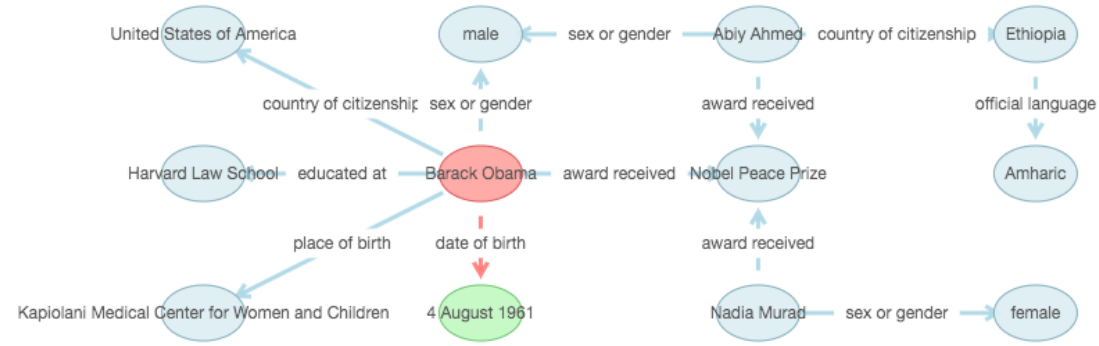
2.


speculative fiction novel
Ron Weasley
spouse
Hermione Granger
sex or gender
female
genre
characters
characters
sex or gender
Harry Potter
characters
Harry Potter
author
J. K. Rowling
date of birth
31 July 1965
place of birth
publisher
country of origin
country of citizenship
place of birth
31 July 1980
Bloomsbury Publishing
country
United Kingdom
country
Yate

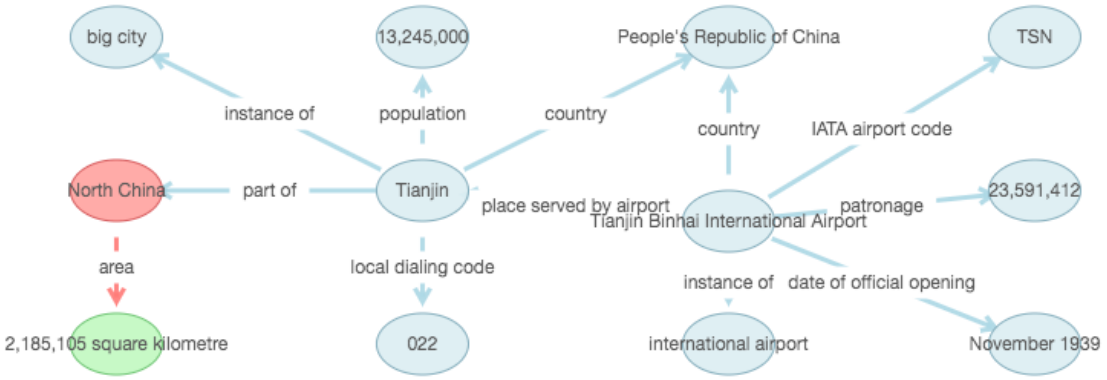
3.


Addison-Wesley
country
Germany
Turing Award
10 January 1938
publisher
award received
date of birth
1969
publication date
The Art of Computer Programming
author
Donald Knuth
place of birth
Milwaukee
main subject
genre
occupation
educated at
algorithm
part of
computer science
field of this occupation
computer scientist
California Institute of Technology

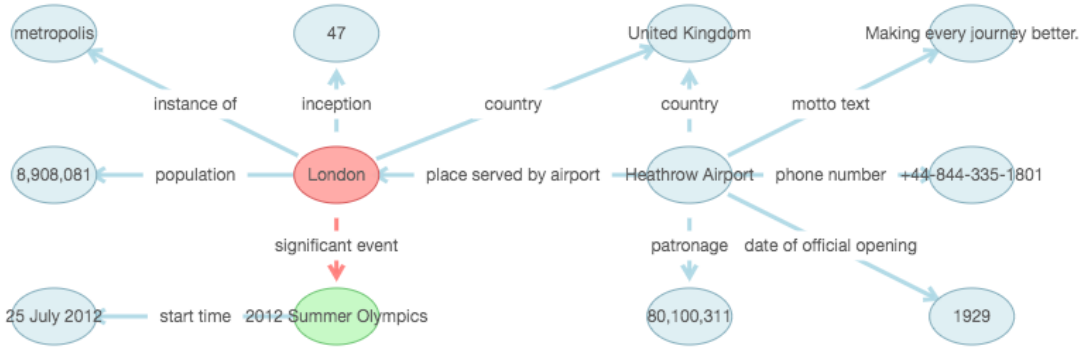
4.


United States of America
male
sex or gender
Abiy Ahmed
country of citizenship
Ethiopia
country of citizenship
sex or gender
award received
official language
Harvard Law School
educated at
Barack Obama
award received
Nobel Peace Prize
Amharic
place of birth
date of birth
award received
Kapiolani Medical Center for Women and Children
4 August 1961
Nadia Murad
sex or gender
female

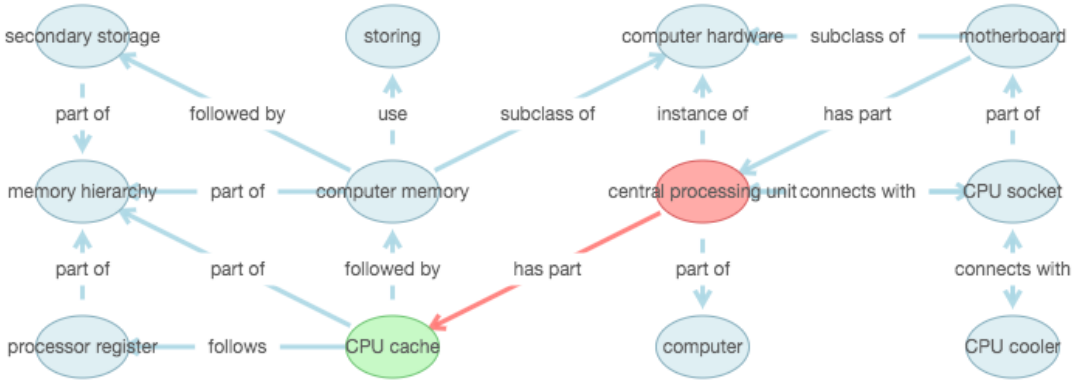
5.


big city
13,245,000
People's Republic of China
TSN
instance of
population
country
country
IATA airport code
North China
part of
Tianjin
place served by airport
Tianjin Binhai International Airport
patronage
23,591,412
area
local dialing code
instance of
date of official opening
2,185,105 square kilometre
022
international airport
November 1939

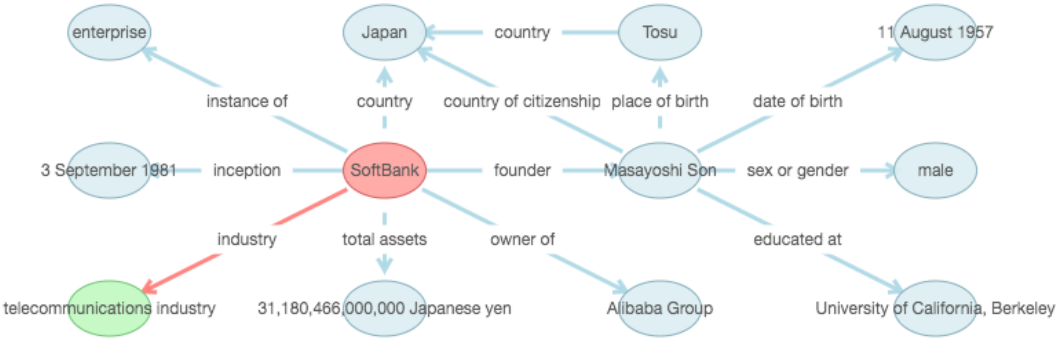
6.


metropolis
47
United Kingdom
Making every journey better.
instance of
inception
country
country
motto text
8,908,081
population
London
place served by airport
Heathrow Airport
phone number
+44-844-335-1801
significant event
patronage
date of official opening
25 July 2012
start time
2012 Summer Olympics
80,100,311
1929

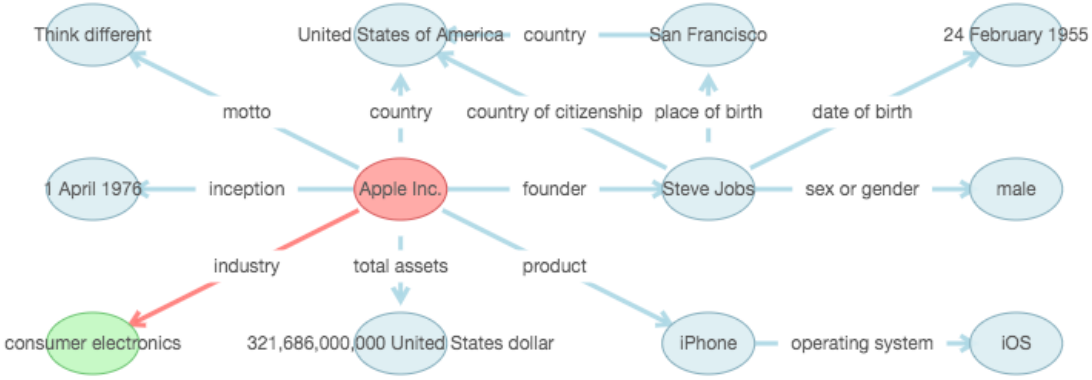
7.


secondary storage
storing
computer hardware
subclass of
motherboard
part of
followed by
use
subclass of
instance of
has part
part of
memory hierarchy
part of
computer memory
central processing unit
connects with
CPU socket
part of
part of
followed by
has part
part of
connects with
processor register
follows
CPU cache
computer
CPU cooler

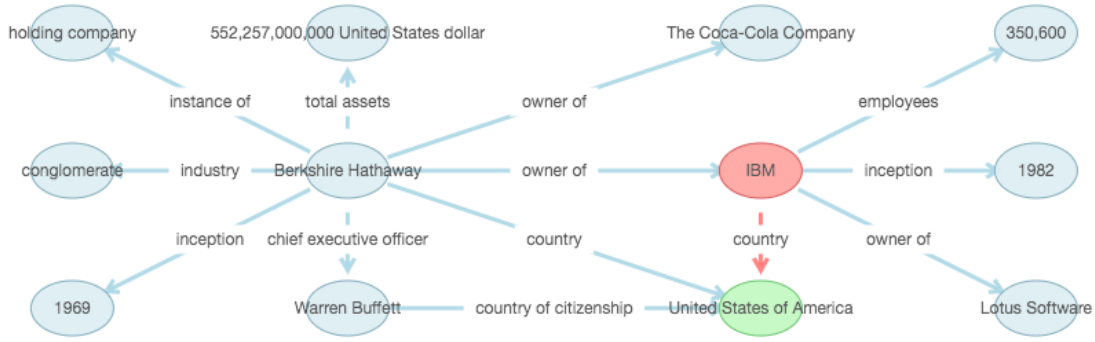
8.


enterprise
Japan
country
Tosu
11 August 1957
instance of
country
country of citizenship
place of birth
date of birth
3 September 1981
inception
SoftBank
founder
Masayoshi Son
sex or gender
male
industry
total assets
owner of
educated at
telecommunications industry
31,180,466,000,000 Japanese yen
Alibaba Group
University of California, Berkeley

9.


Think different
United States of America
country
San Francisco
24 February 1955
motto
country
country of citizenship
place of birth
date of birth
1 April 1976
inception
Apple Inc.
founder
Steve Jobs
sex or gender
male
industry
total assets
product
consumer electronics
321,686,000,000 United States dollar
iPhone
operating system
iOS

10.


holding company
552,257,000,000 United States dollar
The Coca-Cola Company
350,600
instance of
total assets
owner of
employees
conglomerate
industry
Berkshire Hathaway
owner of
IBM
inception
1982
inception
chief executive officer
country
country
owner of
1969
Warren Buffett
country of citizenship
United States of America
Lotus Software