

# 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 Turing 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. Anaphora

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 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).

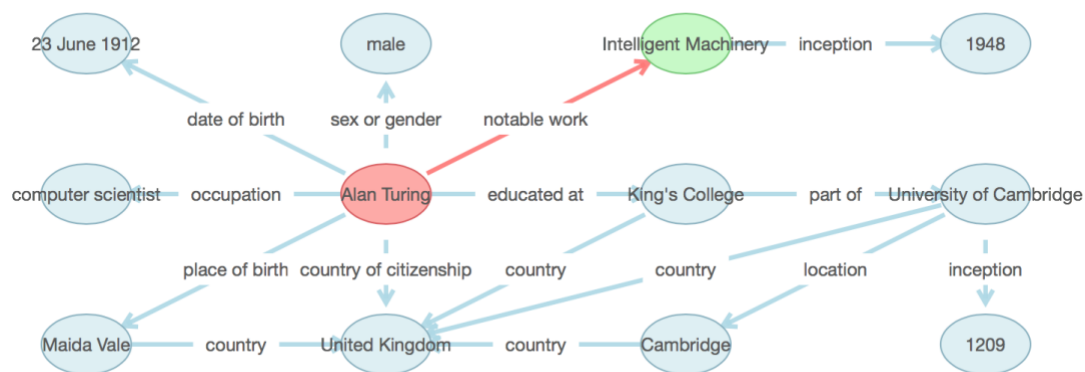
# Labelling:

## 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?

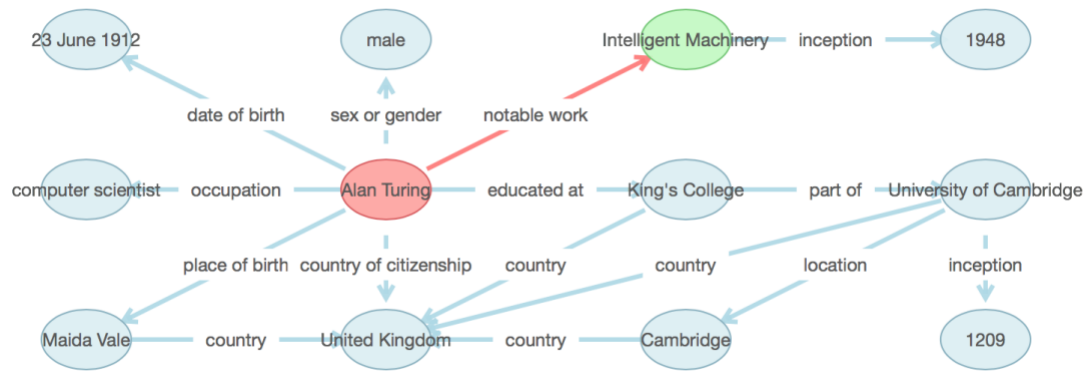
A: Intelligent Machinery! Don't you head of the famous turning test?

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. Turing?

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 inception 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 inception?

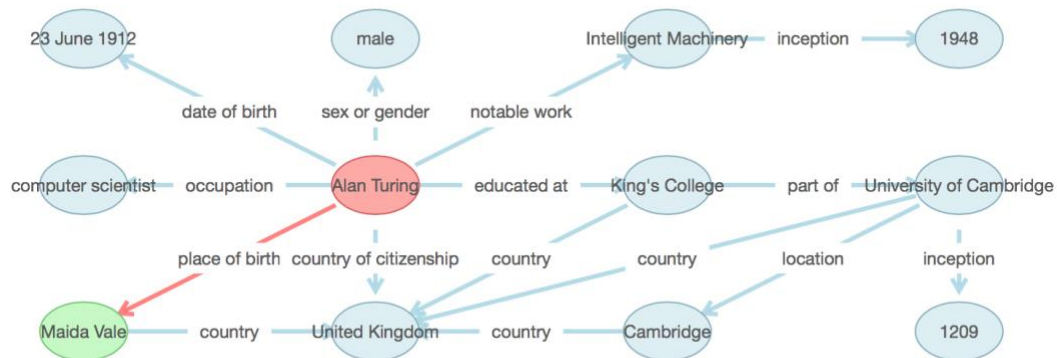
A: 1948.

### 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:



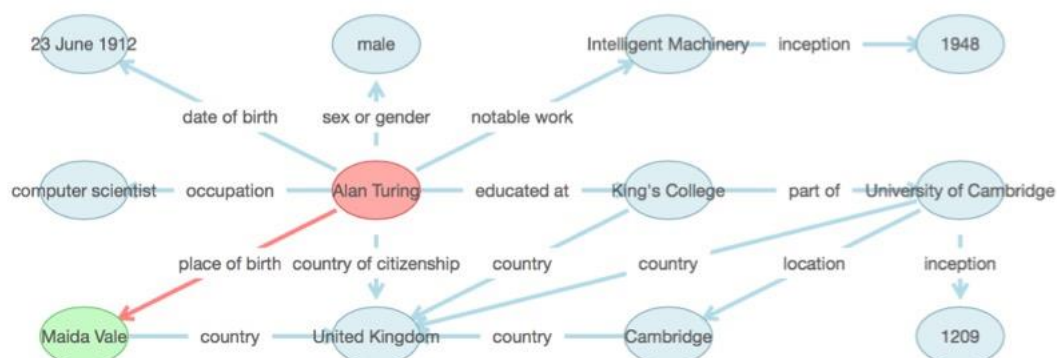
Yes! It is an example of Totally Irrelevant☺:

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?

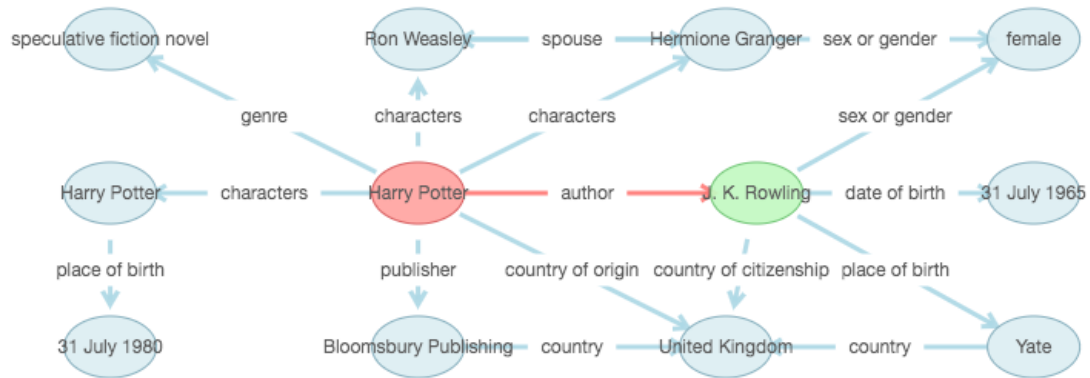
The underlining sentence is considered as totally irrelevant information since it is not directly related to our conversations.

Appendix A: All Grounded Graphs for this Tasks.

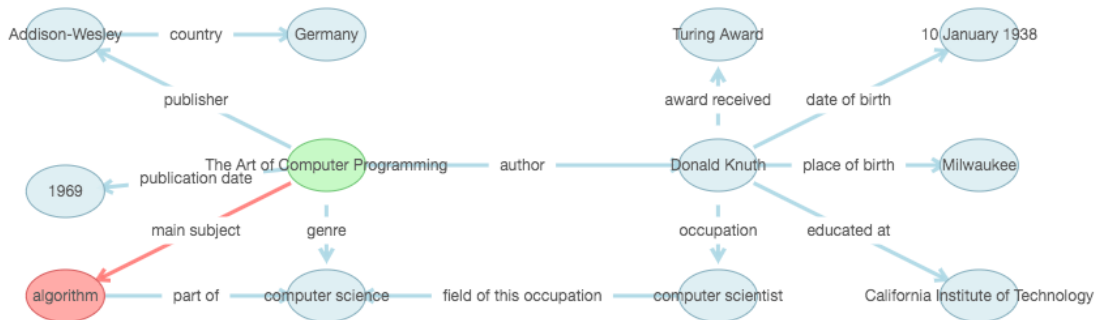
1.



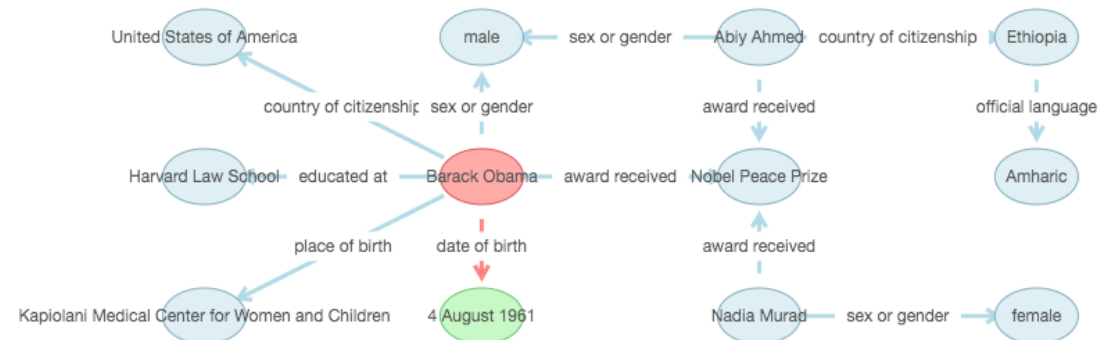
2.



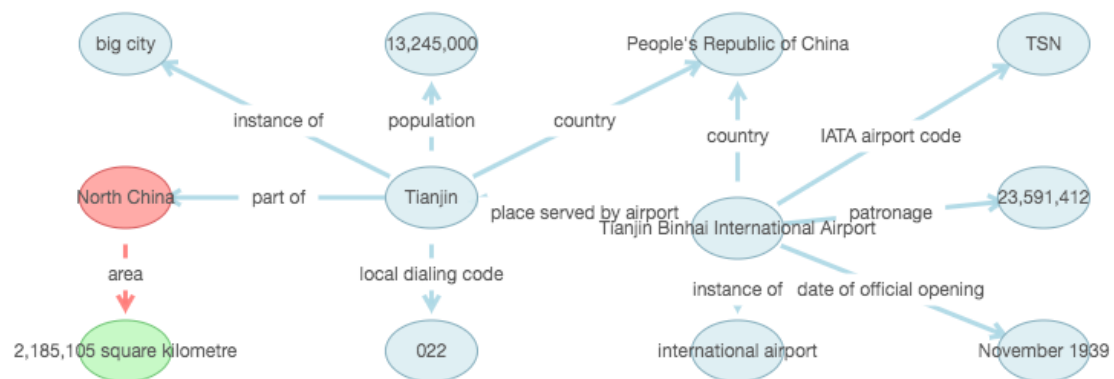
3.



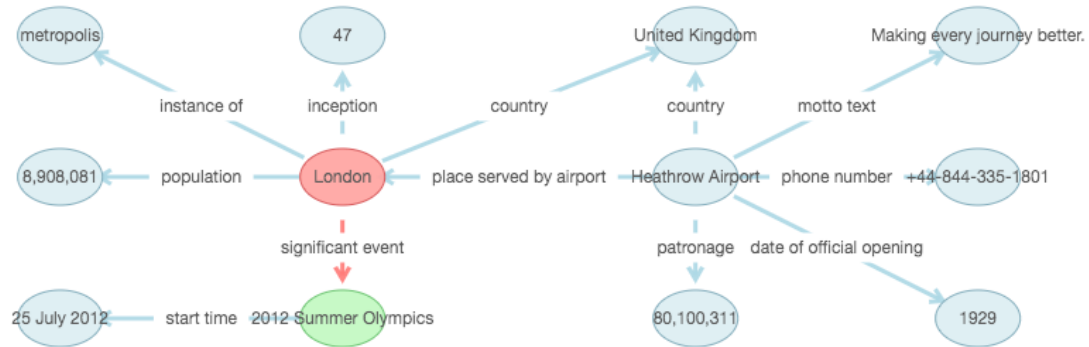
4.



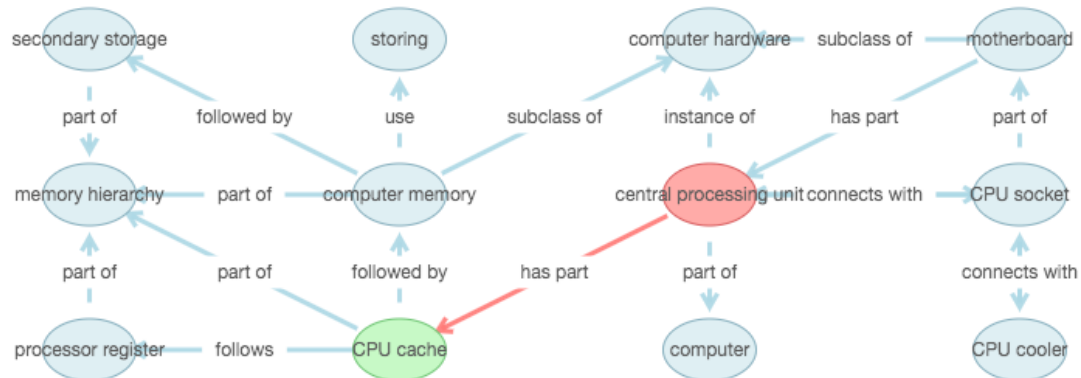
5.



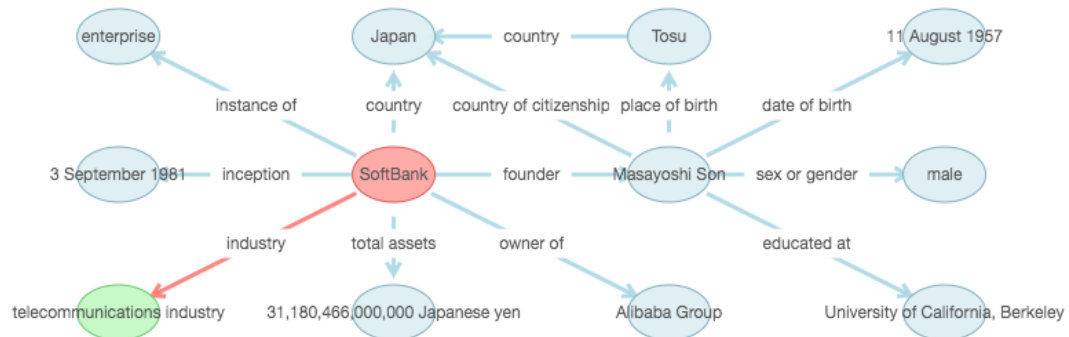
6.



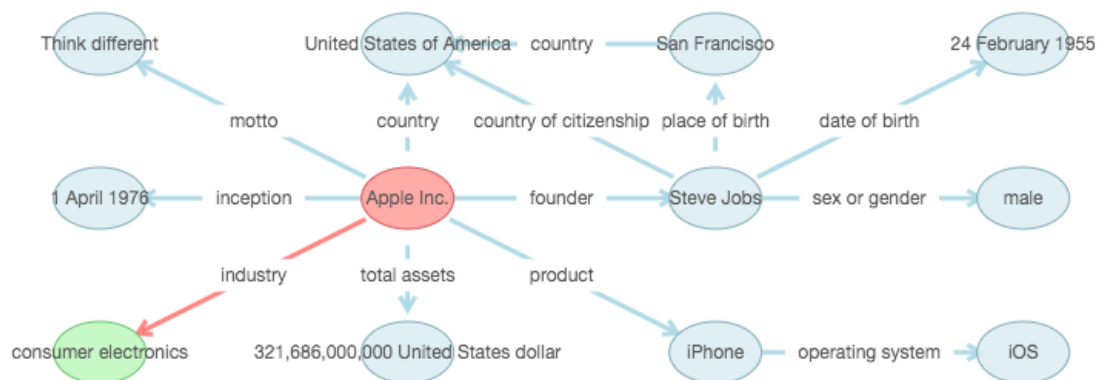
7.



8.



9.



10.

