

Annotation of Relational Triples and Perspectives from User Responses in Context

(Annotation Guidelines)*

Thomas Bellucci
M Artificial Intelligence
t.bellucci@student.vu.nl

Version 1.5

March 2022

Contents

1 Task	3
1.1 Definitions	4
1.1.1 Triples	4
1.1.2 Perspectives	5
2 Data and annotation	6
2.1 Data	6
2.2 Annotation procedure	7
3 Guidelines	9
3.1 Triple annotations	9
3.2 Perspective annotations	13
4 Revisions after Introduction Meeting	14
5 Examples	16
5.1 Example 1	16
5.2 Example 2	16
5.3 Example 3	17
5.4 Example 4	17
5.5 Example 5	18
5.6 Example 6	18

*In partial fulfillment of the requirements for the degree of M Artificial Intelligence at the Vrije Universiteit, Amsterdam

5.7	Example 7	19
5.8	Example 8	19

1 Task

In the context of dialogue systems, the task of Knowledge Extraction (KE) is to extract spans of information from the dialogue in a structured, machine-readable form to allow the acquisition of new world knowledge and establish a personal common ground with a user [3]. For example, given the utterance

(1) *"I went on holiday to Milan last year."*

a KE system could extract the information $\langle user, went\ to, Milan \rangle$ which it could store in a personal database, or triple store. This information can then be reasoned over or be queried by the system later on to respond to questions or generate interesting responses.¹

To represent the information expressed in an utterance, *relational triples* of the form $\langle subject, predicate, object \rangle$ are commonly used; these triples describe two entities (a *subject* and *object*) and a relationship between them (denoted by a *predicate*). The task of the knowledge extraction system is then to mark these triples in the user's utterances.

One difficulty with knowledge extraction in a conversational setting (as opposed to non-conversational text) is that utterances must often be interpreted relative to linguistic context in order to be understood. For example, utterances such as

(2) *"London"* or (3) *"No, believe it was London."*

are commonly spoken in dialogue, yet are uninterpretable without considering the context in which they were said (e.g. *"Did Jim go to Amsterdam?"*). Moreover, as can be seen from example (3), multiple bits of information can be presented within the same utterance and speakers often express a particular *perspective* on the information; in (3), the speaker denies a previous claim (e.g. *"Jim went to Amsterdam"*) and signals a level of (un)certainty.

To accelerate the development of knowledge extraction systems for dialogue capable of capturing information from user input with respect to this *linguistic context*, we will develop a dataset of short dialogues annotated with ground-truth triples and perspectives.² In this document, I will describe the dataset and present guidelines for its annotation.

¹e.g. the user may ask *"Where did I go to again?"* to which the system should reply *"Milan."*

²To our knowledge, no such dataset has to date been developed for English.

1.1 Definitions

In this section, I will informally present to notions of **triple** and **perspective**, which will be used to represent the information asserted by a speaker.

1.1.1 Triples

As stated, we use triples of the form $\langle \textit{subject}, \textit{predicate}, \textit{object} \rangle$ to represent a single, elementary assertion, or *claim*, made by a speaker. *Subjects* and *objects* refer to entities, events, properties or activities; a *predicate* represents a relationship between these. To illustrate,

- (4) *"Mike likes bananas."* (5) *"Jane enjoys walking her dog."*

present each a single claim which we can express as the triples $\langle \textit{Mike}, \textit{likes}, \textit{bananas} \rangle$ and $\langle \textit{Jane}, \textit{enjoys}, \textit{walking her dog} \rangle$.

Subjects, predicates and objects, the *arguments* of the triple, we represent as spans of plain text, e.g. as *"Jane"* and *"likes"*. This makes the definition of triple we use here slightly different from previous work [4], which define a triple as a 3-tuple of arguments, each represented by a so-called Uniform Resource Identifier (URI); we use a plain text representation and consider the assignment of a URI a post-processing step.

Important is that triples contain information on a *single* assertion or claim; to capture several claims, we use several triples. For example, from (6) we infer four triples;

- (6) A.1: *"I am Katana, a nursing student at ODU"*
 B.1: *"Hi Katana. Do you have children?"*
 A.2: *"No, just a pet horse."*
- $\langle \textit{I}, \textit{am}, \textit{Katana} \rangle$
 → $\langle \textit{I}, \textit{am}, \textit{a nursing student at ODU} \rangle$
 → $\langle \textit{you}, \textit{have}, \textit{children} \rangle$ (negated)
 → $\langle \textit{you}, \textit{have}, \textit{a pet horse} \rangle$

In the example above, not all utterances contribute an equal amount of triples; turn A.1 contributes two triples (by virtue of stating two claims about the speaker), while the greeting *"Hi Katana."* in B.1 contributes none (asserting no information).

Also, note that we use *"I"* or *"you"* in the subject position and not *"Katana"* (both pronouns refer to the person *Katana* in this dialogue). We do this as the word *Katana* itself does not serve as a subject in the utterances (it is an object in A.1). We use the subject related to the predicate (in this case *"I"* for the predicate *"am"* and *"you"*

for the predicate *"have"*).

Lastly, note the elliptical construction of *"just a pet horse"* in A.2. On its own, this fragment is hard to interpret as it does not have its own subject and predicate. However, in the context of B.1, this fragment makes sense, providing a contrastive answer to the previous question, namely *"the speaker has instead a pet horse."*. To form a triple for this claim, we inherit the subject and predicate, *"you"* and *"have"*, from the question in B.1.³

1.1.2 Perspectives

As seen in (6), triples provide 'positive' information, even though $\langle \text{you, have, children} \rangle$ is clearly denied by speaker A in the dialogue. To model aspects such as denial or negation (e.g. as signalled by *"No"* in (6)) and uncertainty (e.g. by words such as *"believe"* or *"think"*), we will use **perspectives**. For example, for

(7) *"I think Mike does not like bananas"*

we will use a perspective of the form $\langle \text{polarity, certainty} \rangle$, to mark the polarity and uncertainty of the speaker towards the corresponding claim. A complete annotation of (7), would thus be:⁴

triple: $\langle \text{Mike, like, bananas} \rangle$
 perspective: $\langle \text{not, think} \rangle$

Note however that negation in dialogue must not be signalled explicitly, e.g.

(8) A.1: *"Are you from Amsterdam ?"*
 B.1: *"London ."*

$\langle \text{you, are from, Amsterdam} \rangle$ is implicitly denied and human reasoning is required to arrive at the right interpretation. In this case, we might mark negation by a special [UNK] token appended at the end of each sentence (see guidelines below).

³Note, the tokens *"you"* and *"I"* both refer to speaker A, thus *"you"* can be used instead of *"I"*.

⁴Additional arguments including sentiment may also be included in the perspective. In this work, we will nonetheless consider *polarity* and *certainty* only.

2 Data and annotation

2.1 Data

As no single dataset of dialogue is available for English with sufficient quality and variety to account for the language seen in open-ended social dialogue, data were sampled from three existing dialogue datasets: *PersonaChat* [6], *DailyDialogs* [1] and *Google Circa* [2]. *PersonaChat* is a large corpus of one-on-one, introductory chat conversations. Participants were instructed to get to know each other by conversing about various everyday topics ranging from work and hobbies to family life, taking into account an artificial persona (cf. personality). The dialogues in *PersonaChat* are topically diverse and cover a total of 1000 speaker personas, making them a good test bed for the extraction of personal facts from open-ended social conversations.

The *DailyDialogs corpus* is a collection of short (written) dialogues between two speakers. Unlike *PersonaChat*, the dialogues are situated, that is, the speakers share the environment in which the conversation takes place (e.g. spontaneous conversations on the street, at doctor’s office, and so on). As such, these dialogues may include references to things in the environment, characteristic of social dialogue, but absent from *PersonaChat*.

To conclude the collection, a small number of additional dialogues were sampled from *Google Circa*. *Circa* is a large dataset of polar yes/no questions with direct and indirect user responses, constructed with the intent of training dialogue systems to understand implicit responses, e.g. *"Do you have kids? I got a cat. → [No]"*. The dataset was built to span different social situations and contain a variety of different responses and questions.

Combined, *PersonaChat*, *DailyDialogs* and *Circa* amount to 22.000 dialogues, or approximately 250.000 utterances. As it will not be feasible to annotate the entire dataset within the estimated time frame of the project, we limit the dataset to a random sample of 3600 dialogues, each consisting of three consecutive turns, as per example (6) above.

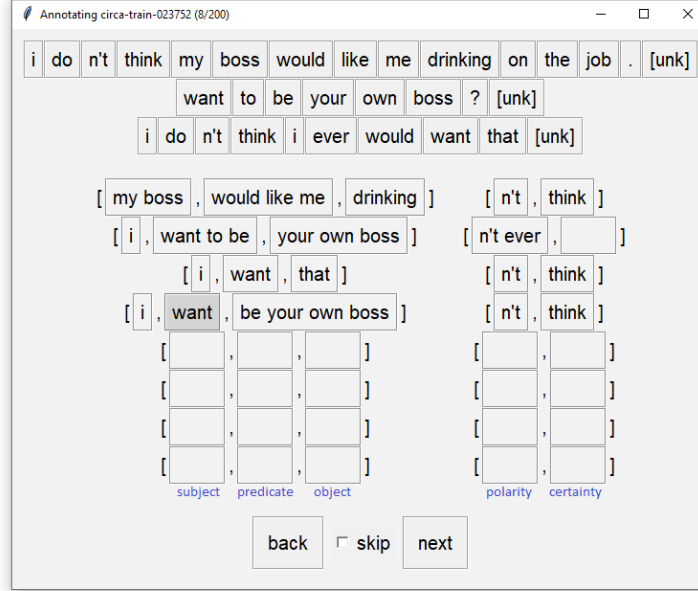


Figure 1: Interface of the annotation tool. Triples and perspectives have been marked. The triple arguments and the polarity and certainty arguments of the perspective are marked in blue.

2.2 Annotation procedure

Each annotator will be provided with a subset of the full dataset, consisting of around 200 short dialogue fragments. To assist in the annotation of these dialogues and ensure a consistent annotation format, a custom annotation tool has been developed.⁵

The interface, shown in Figure 1, shows a list of 8 triples (left column) and their associated perspectives (right column) along with the dialogue (top)⁶. The annotation is performed by assigning tokens from the dialogue to the arguments of the triples and perspectives. To populate the arguments of a triple with tokens from the dialogue, you press the button of the corresponding argument, which will highlight it (like "want" in Figure 1); the user can then click on the corresponding tokens in the dialogue to add them one-by-one to the argument.

The annotation of perspectives with polarity and certainty is performed in the same way as for triples. You click on the corresponding argument followed by the words which signals the polarity or uncertainty. These arguments only have to be marked when the associated triple is negated (often indicated by the words such as *not*, *never*, *no*, *n't*, and so on) or when the speaker raises some doubt (e.g. *maybe*,

⁵<https://github.com/thomas097/Annotation-Tool-for-Master-Thesis>

⁶The number of triples shown is 8 but can be configured in `config.py`.

perhaps, believe, think, might).⁷

After annotating a dialogue, (i.e. populating the arguments of triples (and possibly perspectives) with their corresponding tokens from the dialogue) you can save the annotation by pressing **next**. To skip an annotation (e.g. when the dialogue is unintelligible or its annotation too ambiguous), you press the **skip** checkbox.⁸ If a mistake was made in a previously saved annotation, you can always go back and redo the annotation by pressing **back**. If you only want to redo one argument of a triple, pressing on the argument will reset it.

Lastly, to mark implicit negation (that is, not explicitly signalled by "*not*", "*no*", "*never*", etc.) you can fill in the polarity argument with the [UNK] token.

⁷This was done as positive polarities are rarely expressed overtly, so there are no tokens to be clicked on.

⁸This will make it so that the sample is saved with a special **skipped** flag to indicate a problem.

3 Guidelines

In this section, I will provide a list of annotation guidelines to ensure a consistent annotation result and highlight possible ambiguities and best practices. To clarify these guidelines, I have added a number of examples here; a few more can be found at the end of this document.

3.1 Triple annotations

An **important remark**: The annotation of relational triples in text is inherently ambiguous as triples cannot represent all of natural language (e.g. conditional phrases, adjunctival phrases) and predicate boundaries may be ill-defined.

In case it remains unclear how a dialogue should be annotated given the guidelines below, skip the dialogue using the **skip** button. This will save and mark the annotation with a special flag so it can be looked at later.

- Try to mark triples in all dialogue turns (unless a turn does not express any (relational) information), e.g.

A.1: *"Oh, hi Mark."* → no triple

A.1: *"Hi. I did my homework."* → $\langle I, \textit{did}, \textit{my homework} \rangle$

- **Rule-of-thumb**: Subjects often refer to *entities* or *events*; objects denote *entities*, *activities* or *properties* of entities, and predicates define *relationships* between these. As such, try to keep in mind these types when annotating, e.g.

A.1: *"I did my chores."* → $\langle I, \textit{did}, \textit{my chores} \rangle$

A.2: *"I am very fast."* → $\langle I, \textit{am}, \textit{very fast} \rangle$

A.3: *"I work at the store."* → $\langle I, \textit{work at}, \textit{the store} \rangle$

- **Rule-of-thumb**: keep triples as simple as possible, using the fewest tokens for each argument, e.g. try to leave out **adjunctival prepositional phrases** or **conditionals** that are not essential to the core of the information conveyed;

A.1: *"I worked on my thesis for a little while."*

→ $\langle I, \textit{worked on}, \textit{my thesis} \rangle$

This holds especially for predicates which can balloon because of nested phrases.

- If information is **repeated**, you may mark the triples twice, e.g.

A.1: *"Mike loves cats?"* → $\langle \textit{Mike}, \textit{loves}, \textit{cats} \rangle$ (as confirmed by B.1)

B.1: "Yes, Mike loves cats." $\rightarrow \langle \text{Mike, loves, cats} \rangle$

- In case of **coordination** (*and, or, enumeration*), split into separate triples;

A.1: "What sports do you like?"

B.1: "Soccer, rugby and tennis."

$\rightarrow \langle \text{you, like, soccer} \rangle$

$\rightarrow \langle \text{you, like, rugby} \rangle$

$\rightarrow \langle \text{you, like, tennis} \rangle$

- In case of **ellipsis** (e.g. a dropped subject or predicate), try to complete the corresponding triple with information from the context (e.g. the preceding turn);

A.1: "What sports do you like?"

B.1: "love soccer."

$\rightarrow \langle \text{you, like, soccer} \rangle$

$\rightarrow \langle \text{you, love, soccer} \rangle$

In this example, "you" can be inherited from A.1 as it was spoken by speaker A and thus refers to speaker B.

As shown above, open-questions function as a 'template' to be filled in. Binary yes/no questions imply something to be true (which can then be confirmed or denied by the response). These questions also often contribute a triple, e.g.

A.1: "Did Jim really do that?"

B.1: "Yeah"

$\rightarrow \langle \text{Jim, do, that} \rangle$

- Triple arguments need not be continuous spans in the text, e.g.

A.1: "I can put it away"

$\rightarrow \langle \text{I, can put away, it} \rangle$

but try to keep words within a turn together, e.g.

A.1: "What do you like to do?"

B.1: "eat pizza"

$\rightarrow \langle \text{you, like to do, eat pizza} \rangle$ ("eat" and "pizza" stay together)

- In case of **n-ary relationships** (with roles such as a *giver, receiver* and *thing given*), you may be able to decompose the claim into separate triples;

A.1: "I entertain myself with music."

→ ⟨I, entertain, myself⟩

→ ⟨music, entertain, I⟩

If not, simply ignore the triple (see *Revisions after Introduction Meeting* below).

Similarly, in case of **1-ary relations**, the object should be left blank, e.g.

A.1: "I run."

→ ⟨I, run, ⟩

- In rare occasions, there may not be any subjects mentioned in the dialogue at all. Leave the subjects blank and fill in only the predicates and objects, e.g.

A.1: "went to the cinema." → ⟨ , went to, the cinema⟩

B.1: "saw a good movie?" → ⟨ , saw, a good movie⟩

A.2: "yes, obviously."

Similarly, if a question is asked but no response is given (i.e. no object), you can leave the object argument empty:

A.1: "What sports do you like?"

B.1: "Huh"

→ ⟨you, like, ⟩

- Try to exclude **auxiliary verbs** (e.g. *does*) from the predicate (unless it is the complete predicate), but keep **particles** and **prepositions** (e.g. *to*, *on*) attached to the verb when possible:⁹

A.1: "Where does Mike like to eat at?"

B.1: "An Italian restaurant."

→ ⟨Mike, like to eat at, an Italian restaurant⟩

Particles and prepositions often influence the meaning of the predicate.

- Try to resolve antecedents of referring expressions (pronouns, definite NPs) when possible in addition to the same triple with unresolved arguments;

A.1: "Does Mark like dance music?"

B.1: "Yes, he loves that."

→ ⟨Mark, loves, dance music⟩

→ ⟨he, loves, that⟩

⁹Particles and prepositions are often essential to determine what a predicate means.

Note here that "*Mark*" is a subject, thus can be used to fill in the subject argument, thus can serve as a subject of the other triple.

- If there are few claims made (e.g. only 1 or 2), you can annotate only a subset of triple rows you need and leave the remaining empty.
- The order of the annotations is not important (no need to follow the order triples are presented in the text), but arguments of one triple should be placed together in the same triple row.
- Sometimes *it's* is written as *its*. In this case, you cannot mark the verb *'s* separately (which may be required for, e.g., a predicate). As it will only add noise to the annotation, you may skip this triple.

3.2 Perspective annotations

In the most simple case, a claim is made which has a positive polarity and no uncertainty from the speaker, e.g.

"I am a student" $\rightarrow \langle I, am, a student \rangle$
 "You play fiddle? Yes." $\rightarrow \langle you, play, fiddle \rangle$
 "I want to have a cat." $\rightarrow \langle I, want to have, a cat \rangle$

In this case, there is no need to fill in the perspective column; triples are assumed to have a positive polarity and complete certainty by default.

When the information presented the speaker is denied or the speaker is uncertain about what is said, the perspective column should be used, e.g.

A.1: "My back is hurting."
 B.1: "That is n't good, have you tried a doctor?"
 A.2: "Yes, but I do n't think he listened to me."

From the triples that can be derived, only two of which mark negation or uncertainty:

$\langle My\ back, is, hurting \rangle$ (positive \rightarrow leave blank)
 $\langle That, is, good \rangle$ (negated by *n't* \rightarrow mark *n't*)
 $\langle You, have\ tried, a\ doctor \rangle$ (confirmed by *Yes* \rightarrow leave blank)
 $\langle a\ doctor, listened\ to, me \rangle$ (negated and uncertain \rightarrow mark *n't* and *think*)

For perspectives, we defined the following list of guidelines:

- You only have to mark *polarity* when the triple is negated, e.g. by *not*, *n't*, *never*, *no* or when a previous statement is denied by the other speaker, e.g.

A.1: "Mark loves jazz, right?"
 B.1: "No, not at all."
 $\rightarrow \langle Mark, loves, jazz \rangle, \langle No, \rangle$

- You only have to mark *certainty* when the user indicates to be uncertain about the claim made (e.g. by *believe*, *think*, *might*, *may*, etc.).
- In case of implicit signalling of negation (or uncertainty), mark the polarity (uncertainty) with the [UNK] token. For example,

A.1: "Does he have children?"
 B.1: "I think he only has birds."
 $\rightarrow \langle he, have, children \rangle, \langle '[UNK]', 'think' \rangle$
 $\rightarrow \langle he, has, birds \rangle, \langle , 'think' \rangle$

4 Revisions after Introduction Meeting

During the introduction meeting on April 1st, a number of concerns were raised regarding the definitions and possible ambiguity in the annotations. From this discussion, the following additions were made to the guidelines;

- An utterance can contain **multiple verbs**, e.g. "*I like to watch movies.*" In this case, the boundaries of the predicate become ambiguous;

A.1: "*I — like to watch — movies.*"

A.1: "*I — like — to watch movies.*"

Here, we will follow previous annotation guidelines [5] and place the complexity within the predicate, e.g.

A.1: "*I like to go to the bar.*"

→ $\langle I, \textit{like to go to}, \textit{the bar} \rangle$

We will normalize these predicates at a later stage (for example, to *like_goto*).

- Since it can be difficult to piece apart information about individuals when they are referred to by *their* or *they*, we will ignore expressions that denote **groups**, and will focus on mentions of *single* individuals;

A.1: "*Their names are Ava and John*" → no triple

Nonetheless, it often remains possible to separate members of a group when in a coordinating conjunction. In that case, you may do the following;

A.1: "*John and Alice like trains.*"

→ $\langle \textit{John}, \textit{like}, \textit{trains} \rangle$

→ $\langle \textit{Alice}, \textit{like}, \textit{trains} \rangle$

B.1: "*I like cats and dogs.*"

→ $\langle I, \textit{like}, \textit{cats} \rangle$

→ $\langle I, \textit{like}, \textit{dogs} \rangle$

- A dialogue might contain utterances that are easy to annotate (e.g. "*I love movies*") and ones that are much harder (e.g. conditional phrases; "*If he does that, I will get mad.*"). In this case, annotate the triples you *can* annotate, and ignore assertions you *cannot* (use the *skip* checkbox only if really needed).
- **Back-channeling** by the hearer (e.g. "*ah*", "*uh-huh*", "*wow!*", "*really?*") do not necessarily communicate any information apart from indicating to the speaker that he/she is still listening. These should not be part of a triple.
Also, statements such as **directives** ("*Give me that box.*") don't imply any transfer of information. These should thus also not take part in a triple.

- Although stated in the main guidelines (section 3.1), when the dialogue is ambiguous or incoherent, you can skip the annotation of this dialogue (check the *skip* checkbox).
- It may be possible to infer facts from the dialogue that are not explicitly stated. For example, if the speaker says "*My dogs bit me.*" or "*My two kids love soccer.*", you might infer that the speaker has a dog and two children. This information need not be marked; only information explicitly expressed.

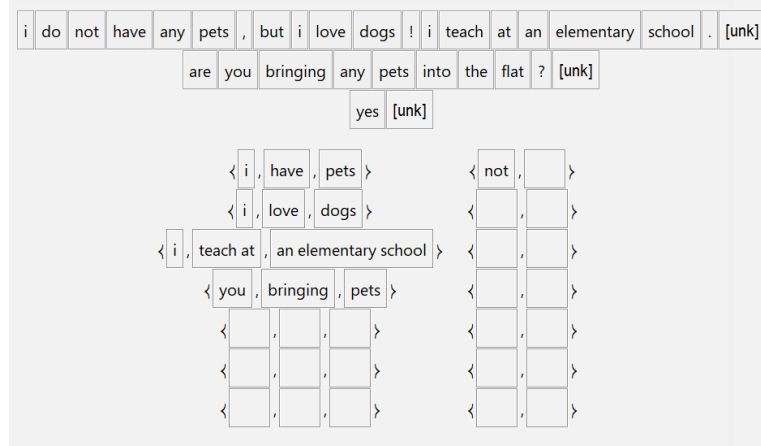
Modifications to annotation software:

In the feedback in the introduction meeting, it was noted that the interface was not very intuitive. The annotation program was revised to include a *skip checkbox* (which can be set to skip a dialogue; see section 2.2). Now, you can freely go back and forth (by pressing the back/next buttons) and the dialogue you skipped will stay skipped (unless the skipped checkbox is unchecked).

5 Examples

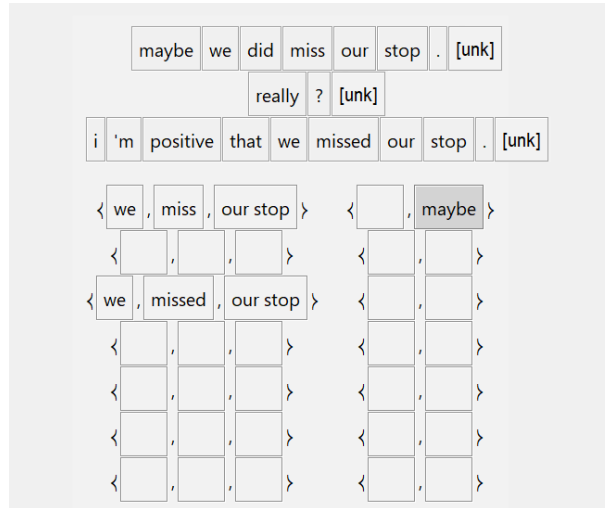
To illustrate how dialogues may be annotated, I have included a few examples below.

5.1 Example 1



* Arguably, the predicate *"bringing"* could be replaced by *"bringing into the flat"*.

5.2 Example 2



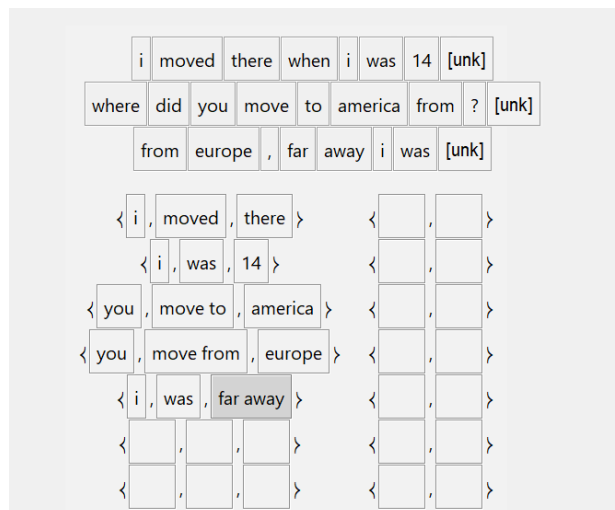
* The middle utterance presents no claim, thus no triple is extracted.

5.3 Example 3

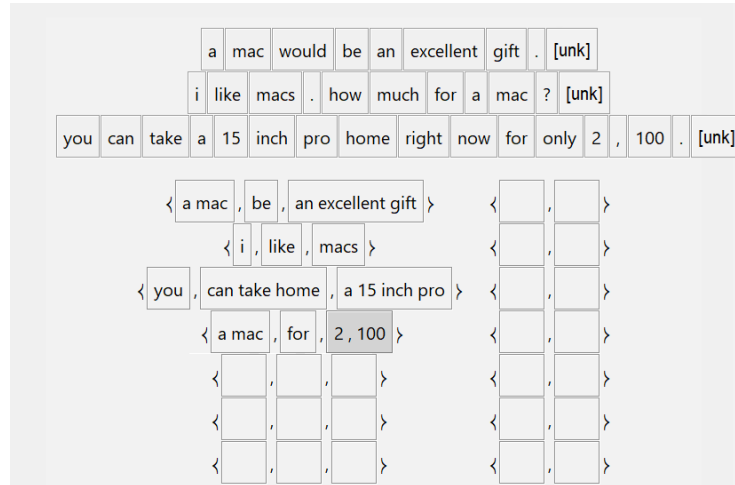


* As no response is given to the first question, the object remains empty. Implicit negation is marked by ”.”.

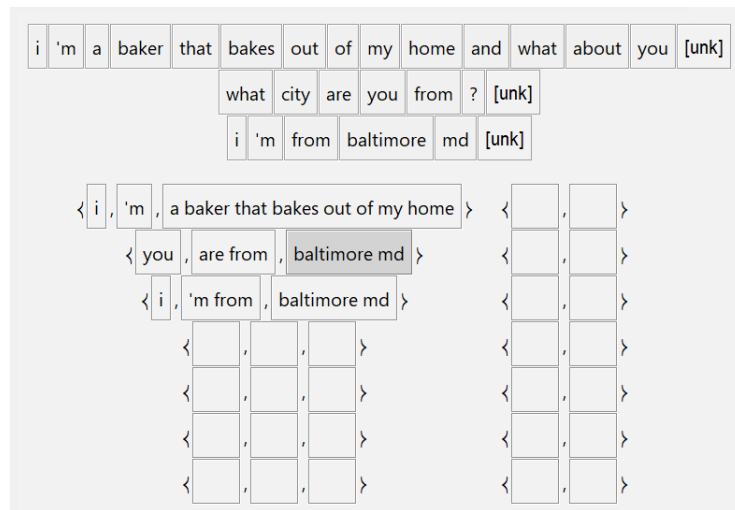
5.4 Example 4



5.5 Example 5

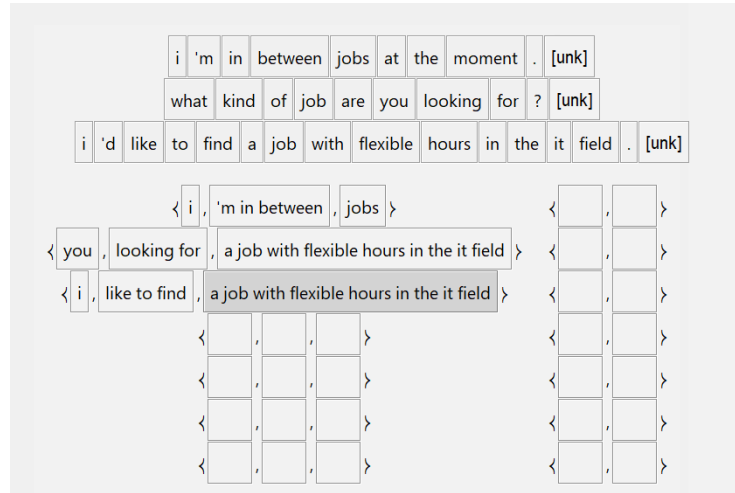


5.6 Example 6



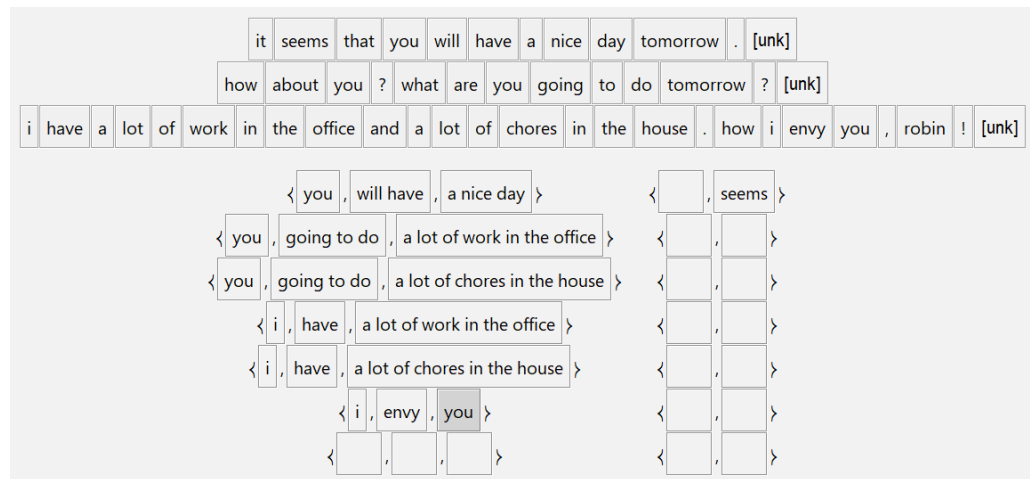
* The first argument "A baker that bakes out of my house" could arguably be reduced to "a baker", however at a loss of specificity.

5.7 Example 7



* "A job with flexible hours in the IT field" may be decomposed into two triples with arguments "a job with flexible hours" and "a job in the IT field", although this would allow for the possibility that the speaker is looking for two jobs.

5.8 Example 8



* In principle, "a lot of chores in the house" and "a lot of work in the office" can be simplified to "a lot of chores" and "a lot of work", respectively.

References

- [1] Yanran Li et al. “Dailydialog: A manually labelled multi-turn dialogue dataset”. In: *arXiv preprint arXiv:1710.03957* (2017).
- [2] Annie Louis, Dan Roth, and Filip Radlinski. “I’d rather just go to bed: Understanding Indirect Answers”. In: *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing*. 2020.
- [3] Brian MacWhinney and William O’Grady. *The handbook of language emergence*. John Wiley & Sons, 2015.
- [4] Piek Vossen et al. “Leolani: A robot that communicates and learns about the shared world”. In: *2019 ISWC Satellite Tracks (Posters and Demonstrations, Industry, and Outrageous Ideas), ISWC 2019-Satellites*. CEUR-WS. 2019, pp. 181–184.
- [5] Sean Welleck et al. “Dialogue Natural Language Inference”. In: *CoRR* abs/1811.00671 (2018). arXiv: 1811.00671. URL: <http://arxiv.org/abs/1811.00671>.
- [6] Saizheng Zhang et al. “Personalizing dialogue agents: I have a dog, do you have pets too?” In: *arXiv preprint arXiv:1801.07243* (2018).