Annotation of Semantic Triples and Perspectives from User Responses in Context

(Annotation Guidelines)*

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

Version 1.1

March 2022

Contents

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

1 Task

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

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

a KE system could extract the information ⟨*user, went to, Milan*⟩ and from it decide to continue the conversation by, e.g., asking *"Where is Milan located?"*.[1]

To represent the information expressed in an utterance, semantic triples of the form ⟨*subject, predicate, object*⟩ are commonly used; these triples describe two entities (the *subject* and *object*) and a relationship between them (the *predicate*). The task of the knowledge extraction system is to mark these triples from the user's utterances in the dialogue.

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

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

are uninterpretable without considering the context in which they were spoken (e.g. *"Where did you go on holiday?"*). Moreover, as can be seen from example (3), multiple bits of information are often presented within the same utterance and speakers often express a particular *perspective* on the information, e.g. denying the information or signalling 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 *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.

2 Definitions

In this section, I will informally present to notions of **triple** and **perspective**, which will be used to represent the assertions made by the user.

[1] In Leolani, these responses are evoked by so-called *thoughts*; reflections on the state of the brain of the robot after receiving new information.

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

Triples of the form ⟨*subject, predicate, object*⟩ are used to represent a single, elementary claim made by a user. The *predicate* represents a relationship between two terms or entities, its so-called *arguments*. For example,

(4) *"Mike likes bananas"*

presents a claim which we can express as the triple ⟨*Mike, likes, bananas*⟩. Most important is that, a triple contains information on a *single* claim made by the user. For example, from (5) we infer four triples (each their own claim);

(5) 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."*

→ ⟨I, am, Katana⟩
→ ⟨I, am, a nursing student at ODU⟩
→ ⟨you, do have, children⟩ (negated)
→ ⟨you, do have, a pet horse⟩

Note the elliptical construction of *"just a pet horse"*. The triple representing this information requires a predicate and subject, *"you"* and *"have"*, which it must inherit from the context (in this case question) in order to be a complete triple.

Moreover, note that triples in (5) are positive, even though ⟨you, have, children ⟩ is denied by the user in the dialogue. To model denial or negation, we will use **perspectives**. As shown by (6), a user may express a perspective on the information presented to signal whether a claim is (believed to be) true or not, or to indicate uncertainty;

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

In order to capture negation and uncertainty, we will use a perspective of the form ⟨*polarity, certainty*⟩, which like triples, mark the polarity and uncertainty of the claim. A complete interpretation of (6), would be:[3]

triple: ⟨*Mike, does like, bananas*⟩
perspective: ⟨*negative, uncertain*⟩

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

3 Data and annotation

3.1 Data

As no single dataset of social dialogue is available with sufficient quality and variety to account for the language seen in open-ended social dialogue, dialogues were sampled from three existing datasets: *PersonaChat* [4], *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 (personality). The dialogues in PersonaChat are topically diverse and cover a total of 1000 speaker personas, making them highly applicable to the setting of open-ended social conversations with different speakers.

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).

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."*. 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 1800 training and 1800 validation (test set) dialogues, each consisting of three consecutive turns; a (user) *context turn*, a (system) *question turn* and a (user) *response turn*, e.g.

A.1: *"I love travelling!"*

B.1: *"Me too! Where did you last go on holiday?"*

A.2: *"Milan."*

3.2 Annotation procedure

Each annotator will be provided with a subset of the full dataset, consisting of 80 (?) turn-triplets. To assist in the annotation process and ensure a consistent annotation format, a custom annotation tool has been developed. The interface, shown in Figure 1, shows a list of 6 triples (left column) and their associated perspectives (right column) along with the dialogue (top). As shown, the annotation is performed by assigning the right tokens to the right triple and perspective arguments.

As explained, triples contain three arguments; a *subject*, a *predicate* and an *object*. To populate the arguments of a triple with tokens from the dialogue, one presses the button of the corresponding argument (e.g. ), which will highlight this argument;

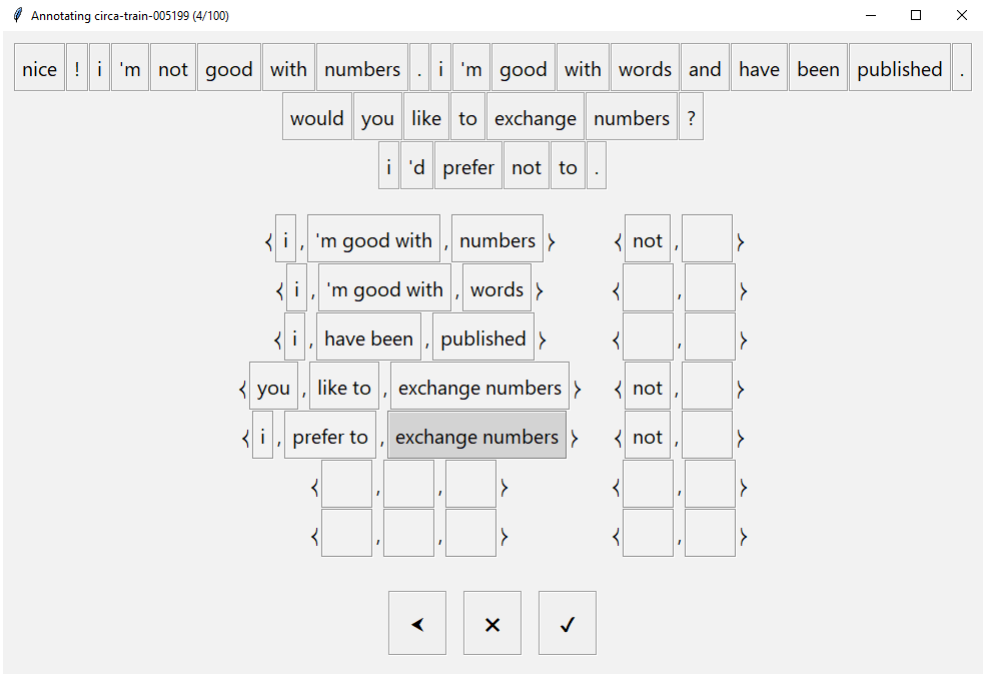


Figure 1: Interface of the annotation tool. Five triples and perspectives have been marked; the remaining two triples were left blank.

the user can then click on the corresponding tokens in the dialogue to fill it in. If a mistake has been made, the user can click again on the argument button which will clear the tokens assigned to that argument.

Like the triples, to annotate the perspectives with polarity and certainty (columns ¬ and ?, respectively), one also clicks on the corresponding argument followed by the token which signals the polarity or certainty. These arguments only have to be marked when the associated triple is negated (often indicated by the tokens *not*, *never*, *n't*, and so on) or when the speaker raises some doubt (e.g. *believe*, *think*).

After annotating the dialogue, that is, populating the triples (and possibly perspectives) with their corresponding arguments, the user can save the annotation by pressing ✓. To skip an annotation (e.g. when the dialogue is unintelligible), you press the ✕ button. If a mistake has been make in a previously submitted annotation, the annotator can go back and redo the annotation by pressing ◂.

4 Guidelines

In this section, I will provide a list of annotation guidelines to ensure a consistent annotation and highlight possible ambiguities and best practices.

4.1 Triple annotation

- Try to annotate triples in all three dialogue turns, irrespective of speaker (system triples can be filtered out when needed).
- Mark triples twice when information is repeated, e.g.

  A.1: *"Mike loves cats?"* → ⟨*Mike, loves, cats*⟩
  B.1: *"Yes, Mike loves cats."* → ⟨*Mike, loves, cats*⟩

- The dialogue might contain information to populate only 1 or 2 triples. In this case, there is no need to fill in all rows; you can leave the remaining rows empty.
- The order of the annotations is not important (no need to follow order of occurrence of triples in the text), but arguments of one triple should be placed in the same row.
- Triple arguments need not be a continuous span in the text, e.g.

  A.1: *"Do you like animals?"*
  B.1: *"Yes."*
  → ⟨*you, do like, animals*⟩

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

  A.1: *"What sports do you like?"*
  B.1: *"Soccer."*
  → ⟨*you, do like, soccer*⟩

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

  A.1: *"What sports do you like?"*
  B.1: *"Soccer, rugby and tennis."*
  → ⟨*you, do like, soccer*⟩
  → ⟨*you, do like, rugby*⟩
  → ⟨*you, do like, tennis*⟩

- In some cases, there may not be a subject mentioned at all in the dialogue (often *I* or *you*). In this case, the subject can be left blank and you can fill in the remainder of the triples, e.g.

A.1: *"went to the cinema."* $\rightarrow \langle$, *went to, the cinema*$\rangle$

B.1: *"saw a good movie?"* $\rightarrow \langle$, *saw, a good movie*$\rangle$

A.2: *"no, not recently."*

Similarly, if a question is asked but no response is given, you can leave the object argument empty:

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

B.1: *"What?"*

$\rightarrow \langle$*you, do like,* $\rangle$

- Try to leave out semantic sugar such as propositional phrases that are not essential to the information conveyed;

A.1: *"I saw a movie on Saturday evening with all my friends."*

$\rightarrow \langle$*I, saw, a movie*$\rangle$

- Include auxiliary verbs (e.g. *do*, *does*) and particles (e.g. *to*, *on*) in the predicate when possible:

A.1: *"What <u>does</u> Mike like <u>to</u> eat?"*

B.1: *"Oven pizza mostly."*

$\rightarrow \langle$*Mike, does like to eat, oven pizza*$\rangle$

- Aim to keep subjects and objects as simple as possible; not all tokens need to be assigned to an argument of a triple (e.g. propositional phrases).

- Try to resolve antecedents of referring expressions (pronouns, definite NPs) when possible; if not, mark pronouns themselves.

A.1: *"Does Mark like bars?"*

B.1: *"Yes, he loves them."*

$\rightarrow \langle$*Mark, loves, bars*$\rangle$

- Lastly, as the examples shown, not all triples require context to be understood. Annotate those triples anyway.

4.2 Perspective annotation

In the most simple case, a claim is made which has a positive polarity and no uncertainty on part of the speaker. For example,

"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." → ⟨*I, want to have, a cat*⟩

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 by the user is denied or the user 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 chiropractor?"*

A.2: *"Yes, but he did n't listen to me about my pain."*

From this example, four triples can be derived, only two of which mark negation or uncertainty:

⟨*My back, is, hurting*⟩	(positive → leave blank)
⟨*That, is, good*⟩	(negated in text → mark *n't*)
⟨*You, have tried, a chiropractor*⟩	(confirmed by *Yes* → leave blank)
⟨*chiropractor, did listen to me about, my pain*⟩	(negative → mark *n't*)

4.2.1 Additional notes

- Only 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 likes jazz, right?"*

  B.1: *"No, not at all."*

  → ⟨*Mark, loves, jazz*⟩ + → ⟨*No,* ⟩

- Only 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 some other token, such as punctuation. For example,

  A.1: *"Does he have children?"*

  B.1: *"I think he only has birds."* (implicitly negates the question)

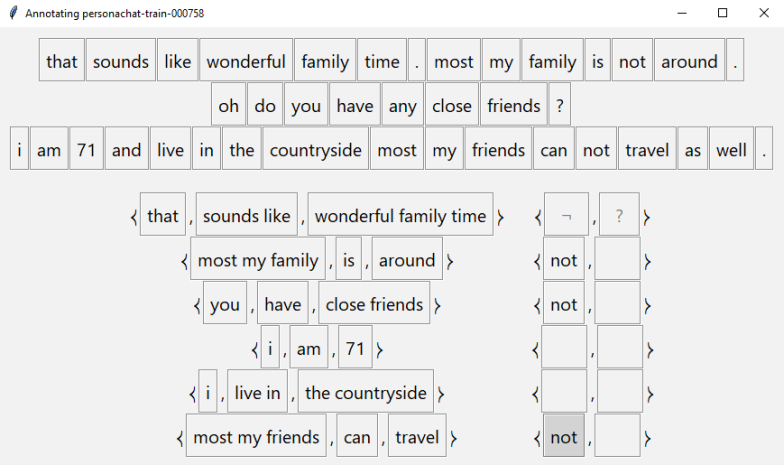
  → ⟨*he, does have, children*⟩ + ⟨*'?',* ⟩

  → ⟨*he, only has, birds*⟩ + ⟨ *, think*⟩

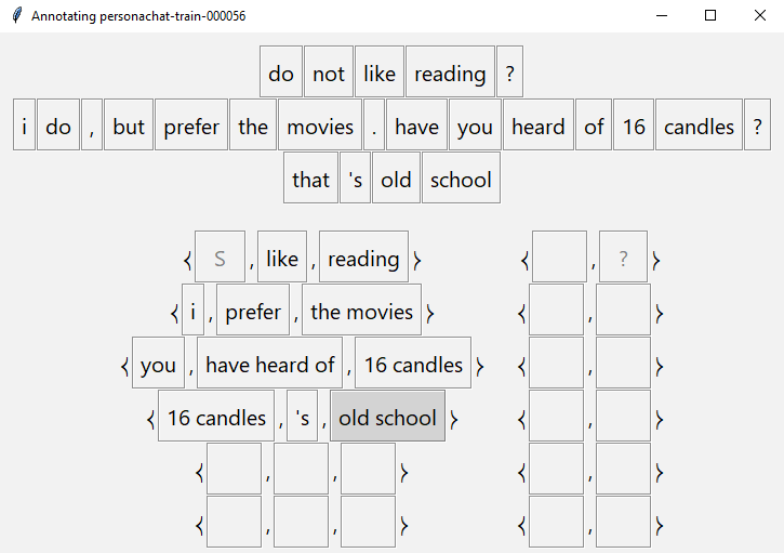
5 Examples

To illustrate how these dialogues ought to be annotated, I have included a couple of additional examples below.

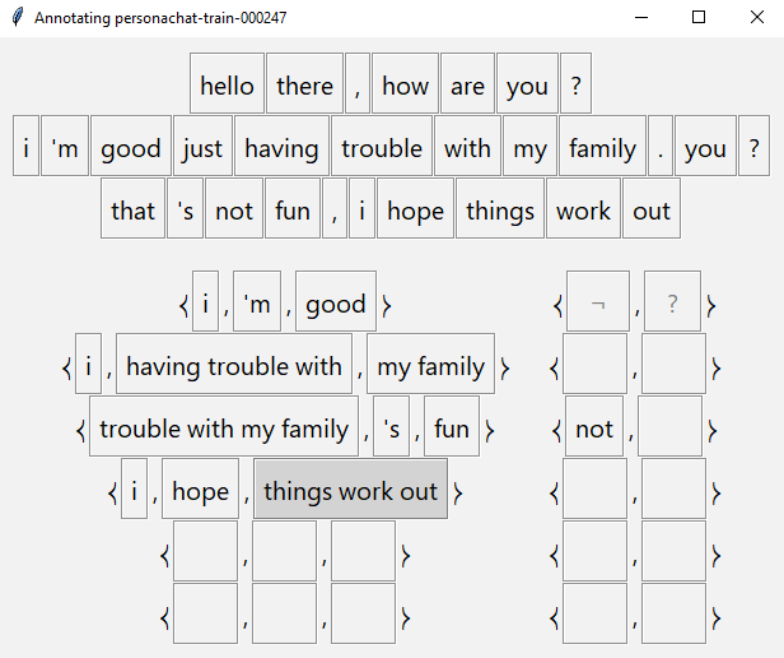
5.1 Example 1



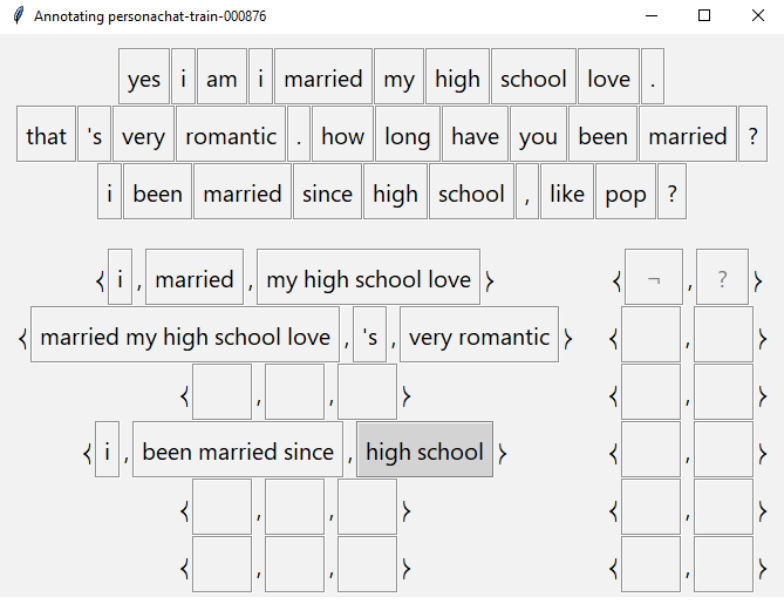
5.2 Example 2



5.3 Example 3



5.4 Example 4



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] Saizheng Zhang et al. "Personalizing dialogue agents: I have a dog, do you have pets too?" In: *arXiv preprint arXiv:1801.07243* (2018).