

Annotation of Relational Triples and Perspectives from User Responses in Context

(Annotation Guidelines)*

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

Version 1.3

March 2022

Contents

1	Task	3
1.1	Definitions	4
2	Data and annotation	5
2.1	Data	5
2.2	Annotation procedure	6
3	Guidelines	8
3.1	Triple annotations	8
3.2	Perspective annotations	11
3.2.1	Additional notes	11
4	Examples	12
4.1	Example 1	12
4.2	Example 2	12
4.3	Example 3	13
4.4	Example 4	13
4.5	Example 5	14
4.6	Example 6	14

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

4.7	Example 7	15
4.8	Example 8	15

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 $\langle user, went\ to, Milan \rangle$ 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, *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 (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 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 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 (i.e. *"Jim went to Amsterdam"*) and signals a level of (un)certainly.

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.

¹In Leolani [4], these responses are evoked by so-called *thoughts*; reflections on the state of the brain of the robot after receiving new information.

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.

Triples of the form $\langle \textit{subject}, \textit{predicate}, \textit{object} \rangle$ are used to represent a single, elementary assertion, or *claim*, made by a speaker. Subjects and objects are used to refer to entities, events, properties or activities; the *predicate* represents a relationship between a these (its so-called *arguments*). To illustrate,

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

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

Important is that a triple contains 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."*
 $\rightarrow \langle \textit{I}, \textit{am}, \textit{Katana} \rangle$
 $\rightarrow \langle \textit{I}, \textit{am}, \textit{a nursing student at ODU} \rangle$
 $\rightarrow \langle \textit{you}, \textit{have}, \textit{children} \rangle$ (negated)
 $\rightarrow \langle \textit{you}, \textit{have}, \textit{a pet horse} \rangle$

Note that not all utterances contribute a triple; turn A.1 contributes two triples (by virtue of stating two claims about speaker A), while the greeting *"Hi Katana."* in B.1 contributes none (asserting no information).

Moreover, note the elliptical construction of *"just a pet horse"* in A.2. On its own, this fragment is hard to interpret. However, in the context of B.1, this fragment makes sense, providing a contrastive answer to the previous question, namely *"Speaker A 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

Lastly, note that triples in (6) are positive, even though $\langle \textit{you}, \textit{have}, \textit{children} \rangle$ 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

³Note, in this dialogue, the tokens *"you"* and *"I"* both refer to speaker A.

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

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

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

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* [5], *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.

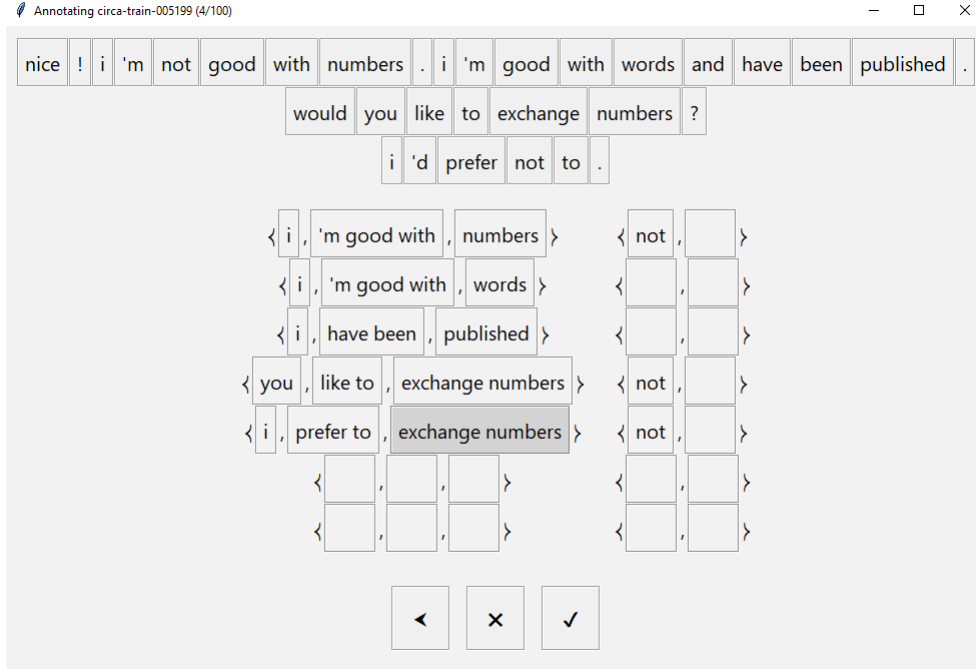


Figure 1: Interface of the annotation tool. Five triples and perspectives have been marked; the remaining two triples were left blank.

2.2 Annotation procedure

Each annotator will be provided with a subset of the full dataset, consisting of around 100 short dialogues. To assist in the annotation of these dialogues and ensure a consistent annotation format, a custom annotation tool has been developed.

The interface, shown in Figure 1, shows a list of 7 triples (left column) and their associated perspectives (right column) along with the dialogue (top)⁵. 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 "Exchange numbers" 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 the same way as for triples. You click on the corresponding argument followed by the token which

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

⁵The number of triples is configurable.

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. *believe*, *think*).⁶

After annotating a dialogue, that is, populating the arguments of triples (and possibly perspectives) with their corresponding tokens, you can save the annotation by pressing ✓. To skip an annotation (e.g. when the dialogue is unintelligible or its annotation ambiguous), you press the × button.⁷ If a mistake was made in a previously submitted annotation, you can always go back and redo the annotation by pressing ◀. If you only want to adjust one argument, pressing on the argument will reset it.

⁶Done as positive polarities are rarely overtly expressed, so there are no tokens to be clicked on.

⁷This will save the sample 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 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) and predicate boundaries may be ill-defined. These triples thus only provide a shallow understanding of meaning.

In case it remains unclear how a triple should be annotated given the guidelines below, skip the dialogue using the \times button. This will save and mark the annotation with a special **skipped** 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.*" $\rightarrow \emptyset$

while,

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

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

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

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

- **Rule-of-thumb**: keep triples as simple as possible, using the fewest tokens for each argument, e.g. try to leave out **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.*"

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

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

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

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

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

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

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

→ $\langle \textit{you}, \textit{like}, \textit{soccer} \rangle$

→ $\langle \textit{you}, \textit{like}, \textit{rugby} \rangle$

→ $\langle \textit{you}, \textit{like}, \textit{tennis} \rangle$

- 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."*

→ $\langle \textit{you}, \textit{like}, \textit{soccer} \rangle$

→ $\langle \textit{you}, \textit{love}, \textit{soccer} \rangle$

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"*

→ $\langle \textit{Jim}, \textit{do}, \textit{that} \rangle$

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

A.1: *"I can put it away"*

→ $\langle \textit{I}, \textit{can put away}, \textit{it} \rangle$

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

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

B.1: *"eat pizza"*

→ $\langle \textit{you}, \textit{like to do}, \textit{eat pizza} \rangle$ ("*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."*

→ $\langle \textit{I}, \textit{entertain}, \textit{myself} \rangle$

→ $\langle \textit{music}, \textit{entertain}, \textit{I} \rangle$

- 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."* → $\langle \text{ , went to, the cinema} \rangle$

B.1: *"saw a good movie?"* → $\langle \text{ , saw, a good movie} \rangle$

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"*

→ $\langle \text{you, like, } \rangle$

- Try to exclude **auxiliary verbs** (e.g. *does*) from the predicate, but keep **particles** and **prepositions** (e.g. *to*, *on*) attached to the verb when possible:⁸

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

B.1: *"An Italian restaurant."*

→ $\langle \text{Mike, like to eat at, an Italian restaurant} \rangle$

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

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

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

→ $\langle \text{Mark, loves, dance music} \rangle$

→ $\langle \text{he, loves, that} \rangle$

- 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.

⁸Particles and prepositions are often essential to determine what a predicate means.

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

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:

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

3.2.1 Additional notes

- 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 likes jazz, right?"
 B.1: "No, not at all."
 $\rightarrow \langle \text{Mark}, \text{loves}, \text{jazz} \rangle, \langle \text{No}, \text{ } \rangle$

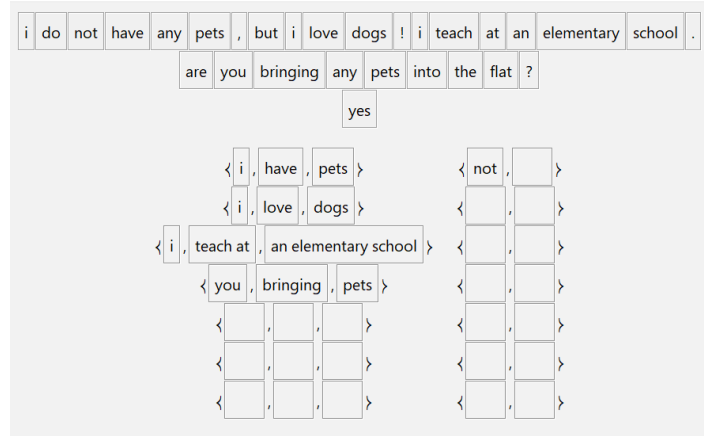
- You only have to mark *certainty* when the user indicates to be uncertain about the claim made (e.g. by *believe*, *think*, *might*, *may*, etc.).
- In case of implicit signalling of negation (or uncertainty), mark the polarity (uncertainty) with some other token, such as punctuation. For example,

A.1: "Does he have children?"
 B.1: "I think he only has birds."
 $\rightarrow \langle \text{he}, \text{does have}, \text{children} \rangle, \langle '?', \text{'think'} \rangle$
 $\rightarrow \langle \text{he}, \text{only has}, \text{birds} \rangle, \langle \text{' ', 'think'} \rangle$

4 Examples

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

4.1 Example 1



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

4.2 Example 2



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

4.3 Example 3

where	is	your	family	at	
is	your	family	doing	okay	?
my	mom	has	been	sick	again
.					

<	your family	,	is at	,		>	<		,		>
<	your family	,	doing	,	okay	>	<	.	,		>
<	my mom	,	been	,	sick	>	<		,		>
<		,		,		>	<		,		>
<		,		,		>	<		,		>
<		,		,		>	<		,		>
<		,		,		>	<		,		>

<
×
✓

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

4.4 Example 4

i	moved	there	when	i	was	14	
where	did	you	move	to	america	from	?
from	europe	,	far	away	i	was	

<	i	,	moved	,	there	>	<		,		>
<	i	,	was	,	14	>	<		,		>
<	you	,	move to	,	america	>	<		,		>
<	you	,	move from	,	europe	>	<		,		>
<	i	,	was	,	far away	>	<		,		>
<		,		,		>	<		,		>
<		,		,		>	<		,		>

<
×
✓

4.5 Example 5

[illegible]

4.6 Example 6

i 'm a baker that bakes out of my home and what about you

what city are you from ?

i 'm from baltimore md

< i , 'm , a baker that bakes out of my home >

< you , are from , baltimore md >

< i , 'm from , baltimore md >

< , , >
 < , , >
 < , , >
 < , , >
 < , , >

< , , >
 < , , >
 < , , >
 < , , >
 < , , >

< , , >
 < , , >
 < , , >
 < , , >
 < , , >

< , , >
 < , , >
 < , , >
 < , , >
 < , , >

< , , >
 < , , >
 < , , >
 < , , >
 < , , >

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

4.7 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.

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