Annotation Task:
Clarification Requests Regarding Weak vs. Strong Uptake

julian.schloeder@gmail.com

Please read these instruction carefully before starting the annotation.
If something is not clear, please do not hesitate to ask us. Thanks for your participation!

You are asked to annotate possible clarification requests (CRs) in dialogue excerpts extracted from the AMI meeting corpus. In the dialogues, a group of four people pretend to be the design team of a remote control (for a television): they are provided with this goal and then converse/discuss freely. The corpus has been annotated with dialogue acts. The excerpts you are given contains a full turn by one person P, a question directed at P, and P's answer to that. The question is marked with an arrow `-->`. Note that even though some of these may not be grammatical questions, they have previously been annotated as questions. You will find examples from the corpus below.

Clarification requests indicate problems with earlier utterances. These problems can be at different levels: for instance, the utterer of the CR may not have heard the speaker, or she may not know the meaning of a word used. Consider the following examples; the first clarifies word-meaning, the second one an acoustic problem.

```
D: So , I'm pretty confident that our remote control fits into the furniture category .
D: And also um the feel of material this year um is expected to be spongy .
A: Okay .
D: Um so hopefully our remote control reflects that s in some way .
--> B: What does it mean , spongy ?
D: Uh sort of um squishy .
```

```
A: Mm . Well , we're under cost then .
A: Alright .
--> C: We're over ?
A: No , we're under .
```

Some of the questions are CRs, others might not. The following example is not a clarification question; the asker is requesting novel information, moving the dialogue forward.

```
C: What ?
--> A: Your surname.
C: Uh Abbing . A_ B_ B_ I_ N_ G_ .
```

The main interest of this study are issues at the *intensional* level, for instance, with recognising the intention/goal of the speaker (what the intended function of the utterance is) or with figuring out how to fulfil the intention/goal of the speaker (adopting a belief or carrying out an instruction). We call this level **uptake**.

We ask you to classify the questions as either **not CR**, **low level CR**, **weak uptake CR** or **strong uptake CR**. Here is an overview of the full annotation scheme. Please read carefully the explanations for each class on the next page.

- `not`
- `low`
- `weak`
- `strong`
- `ambig-cr`
- `ambig-level`
- `ambig-uptake`

not Clarification: annotate a question as '`not`' if it does *not* serve to better the askers understanding of the previous utterance. In the example below, `C`'s question (while prompted by `A`) is not clarifying `A`'s utterance.

```
A: By next meeting , okay that will be great .
A: Uh Okay ,
A: so there's any questions or uh first of all about uh this project ?
--> C: What is the goal of the project ?
A: Uh the goal of the project I think maybe I'll uh hand out to the Ed , okay ,
not
```

low-level CR annotate a question as '`low`' if it indicates that the asker has not fully understood the *semantic / propositional content* of the previous utterance. This includes, for example, *acoustic problems*, *word meaning* (see above) or *reference resolution* as in the next example:

```
B: Um , it didn't work anymore .
--> D: The laptop ?
B: The entire Windows uh
low
```

Weak uptake: annotate a CR as '`weak`' if it indicates semantic understanding, but that the CR utterer *has not fully understood* the speaker's goal (the intended function of their utterance). The prototypical case is *speech act determination* as in the example below.

```
C: and we are going to need , um , a plan for that
--> A: do you want me to make one ?
C: yes , if you don't mind
weak
```

Strong uptake: annotate a CR as '`strong`' if it indicates the CR utterer *has understood* the speaker's goal (their intention), *but does not yet ground it because he wants more information.* Typical examples are *why*-and *how*-questions (but don't just blindly tag every why and how as strong uptake!), as well as *contradicting belief.*

```
B: we should go with bluetooth
--> A: are you sure ?
B: um , I guess
strong
```

Ambiguous: Sometimes it may not be possible to decide whether a CR corresponds to low-level, weak or strong uptake, maybe due to a lack of context. In those cases, annotate the CR as one of the ambiguity classes. In the following example, it is unclear what A is responding to exactly, so it is hard to determine what they are clarifying. Please don't overuse this category: whenever possible, do make up your mind and annotate each CR as either '`low`' '`weak`' or '`strong`'.

```
D: Yeah .
--> A: and the other stuff
D: And the speech recognition ,
ambig-cr
```

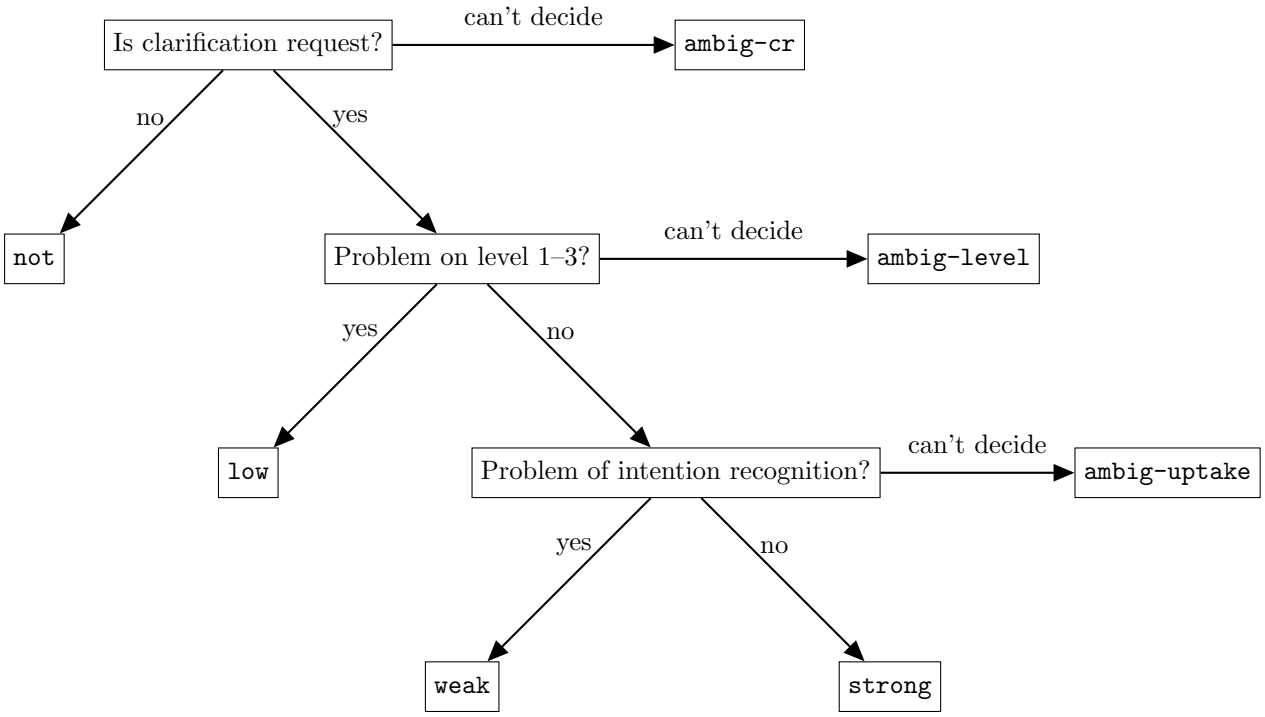
On the next page, you will find a decision tree that will guide your annotation process. Please enter your annotation in one line directly below the 3-turn dialogue following this scheme: {not, low, weak, strong, ambig-cr, ambig-level, ambig-uptake}

Examples:

```
K: 1st context turn
K: 1st context turn
K: 1st context turn
--> I: QUESTION
K: 2nd context turn
ambig-level

K: 1st context turn
--> I: QUESTION
K: 2nd context turn
weak
```

The following decision tree indicates in which order and manner you should reach your decision. Please follow it closely.



Completed annotations should be send to Julian; `julian.schloeder@gmail.com`