

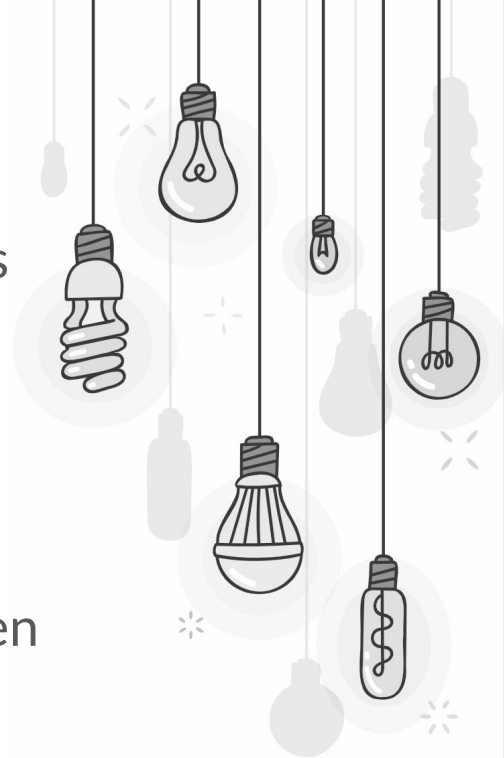
Clarifying implicit and underspecified Phrases in Instructional Text

Talita Anthonio*, Anna Sauer* and Michael Roth

LREC 2022, Marseille

Motivation & background

- * Language contains *implicit* and *underspecified* elements
 - * **Implicit references:** Call ___ for an appointment.
 - * **Fused-heads:** It's about 5 ___.
- * We use such language based on the assumption that people can derive the meaning from the **context**
- * Inferring **the missing elements** can be challenging when there are multiple plausible interpretations of ____
→ humans and computers
- * But the notion that multiple continuations/fillers are plausible is generally neglected in existing **NLP** tasks



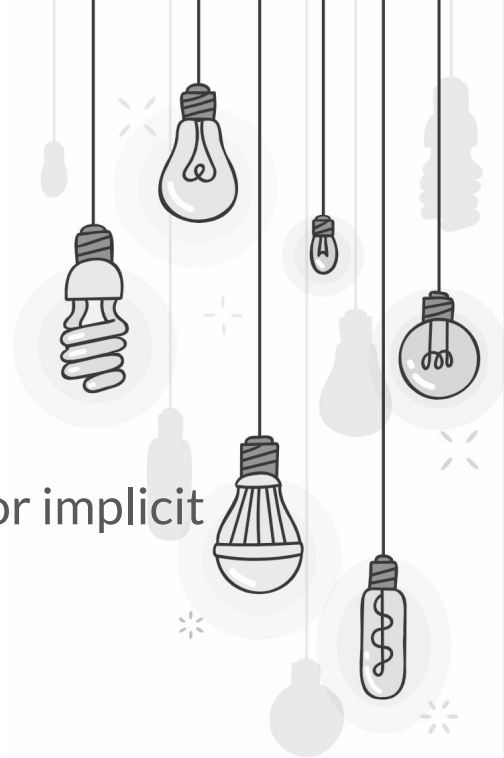
Approach

- * We construct a **dataset** of implausible/plausible clarifications of implicit/underspecified phrases automatically using **revision histories** from wikiHow
- * Previous work (Antonio et al., 2020) showed that revisions in **wikiHow** can have clarifying functions for implicit and underspecified elements

original sentence: Leave ___ out in the sun

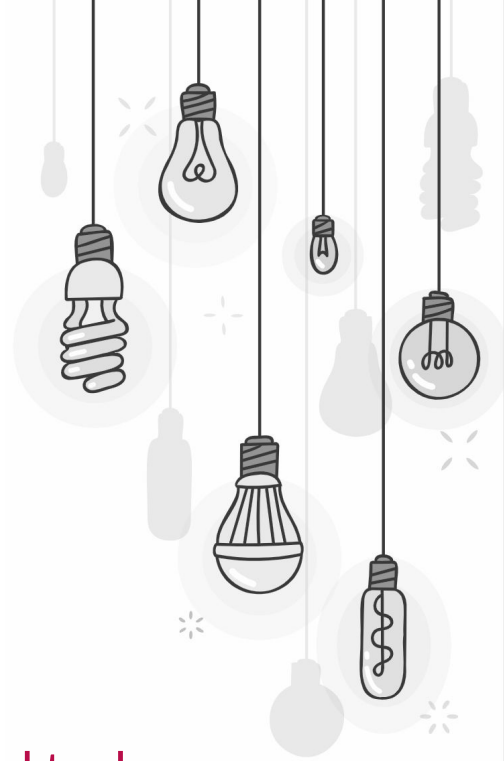
revised sentence: Leave **the vacuum** out in the sun

- * We generate additional *fillers* (henceforth **clarifications**)
- * We determine the **plausibility** of different clarifications for a given sentence in their context via **annotations**



Contributions

1. We release **CLAIRE**: a dataset of sentences with implausible and plausible clarifications of implicit and underspecified phrases in instructions
2. We introduce the task of **distinguishing** between plausible and implausible clarifications for implicit and underspecified phrases
3. We provide several **baseline** models for our introduced task
4. We take a closer look at **conflicting interpretations** of implicit/underspecified phrases



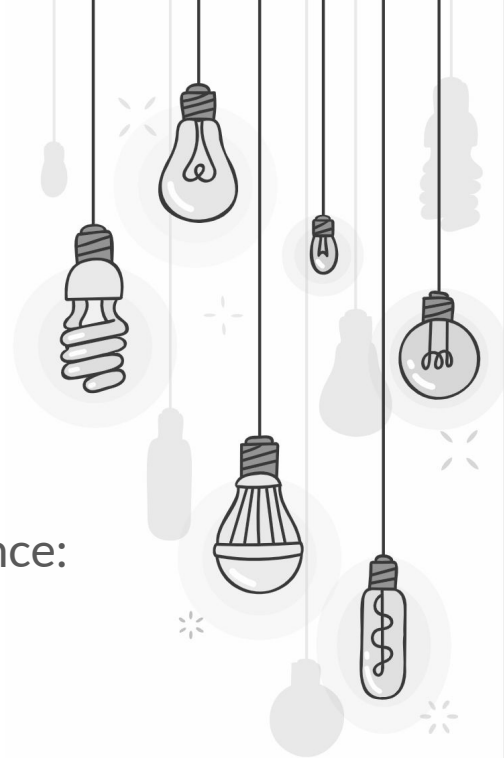
Data creation

How can we create a dataset with implicit and underspecified phrases and their clarifications from Wikihow revisions?

Data creation I

Step 1: extract subsets from wikiHowToImprove (Antonio et al., 2020)

- * We take the sentences where the edit contains *an insertion of a contiguous phrase which allows us to find*
 - * **implicit/underspecified** element in the original sentence:
original: Call ____ for an appointment
 - * which is clarified in the revised sentence:
revised: Call **the salon** for an appointment
- * We use **heuristics** to automatically select insertions that resolve four phenomena of implicit/underspecified language



Phenomena

Phenomena	Example	Train	Dev/Test
Implicit references	Rinse ___ before assembling	996	125
Fused-head	Some ___ like tricks, some like races	999	125
Noun compound	Line a large baking sheet with ___ foil	1000	125
Metonymy	The blanket should be snug around your baby's _____ but not tight	1000	125
total		3395	500

Table 1: Overview of the four phenomena and their frequency distribution in CLAIRE.

Data creation II

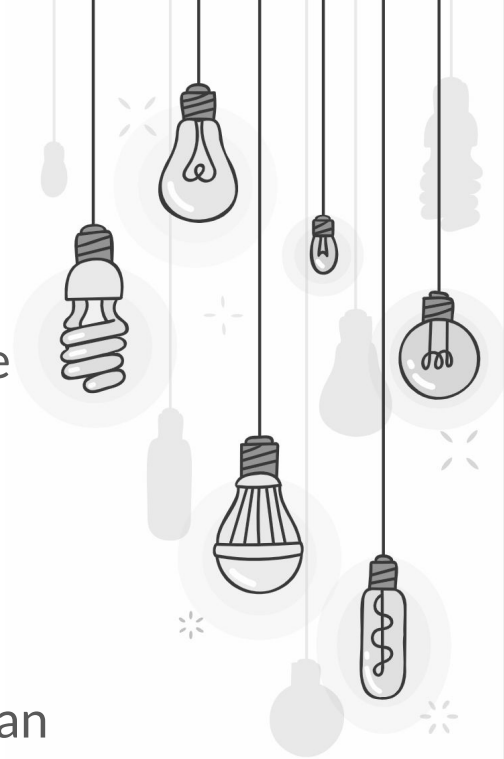
Step 2: generate additional clarifications

- * We use **transformer-based language** models to generate additional clarifications for the sentence in **its context**

sentence: perfect your ____ posture

clarifications: *body, walking, core, gym, target*

- * Previous work showed that transformer-based models can be used to generate the **human insertion** (Antonio & Roth, 2021)



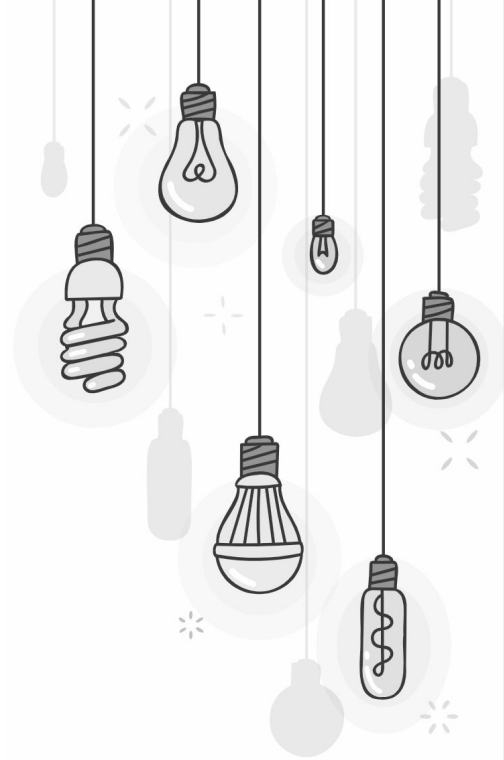
Data creation III

Step 3: select four clarifications that are semantically diverse from each other

- * We select four generated clarifications using *k*-means

Step 4: collect annotations

- * We collect annotations for all **five clarifications** (*generated + human-insertion*) to determine their plausibility



Annotation set-up

Task

- * HIT on Amazon Mechanical Turk for train, dev and test
- * 4 judgements per instance (*sentence + clarification*)
- * rate the plausibility of the clarified part in the given context (1-5)

Qualifications

- * HIT approval rate of 95 %
- * Location should be US/UK
- * Qualification test: 4/4 questions correct

Annotation set-up

Interface

Read the text below and indicate if the underlined part makes sense in the given how-to guide.

Text

How to Braai Steak

Part Three: Braai the Steak

(...)

4. Expose each side to the flame.

(...)

The exact timing can vary depending on the conditions of your grill, though.

5. Cook your steaks to medium rare.

Many braai purists insist that steaks must be cooked to medium rare.

Question

On a scale from 1 to 5, does the underlined part make sense in the given how-to guide?

(1=complete nonsense, 5=definitely makes sense; ratings of 0 will be rejected)



From ratings to classes

- * **Implausible** →
avg. rating ≤ 2.5
- * **Plausible** →
avg. rating ≥ 4.0
- * **Neutral** →
avg. rating > 2.5 & < 4.0

Data description

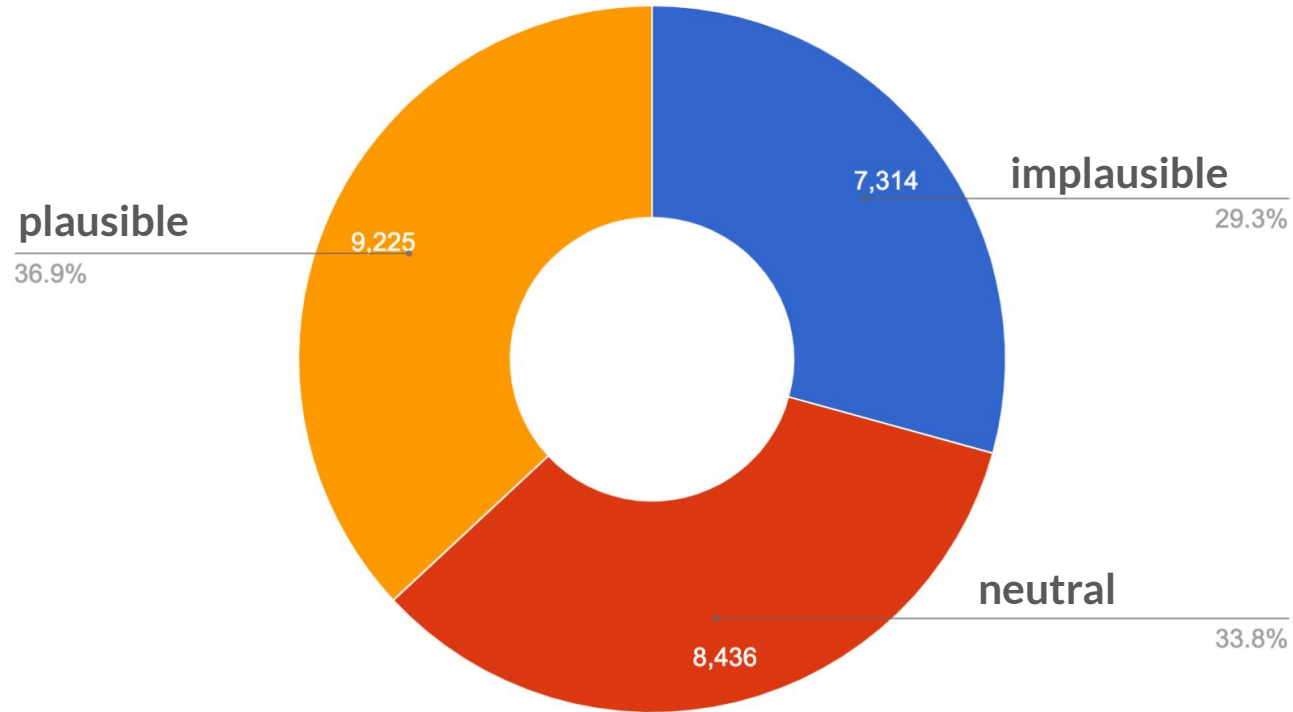


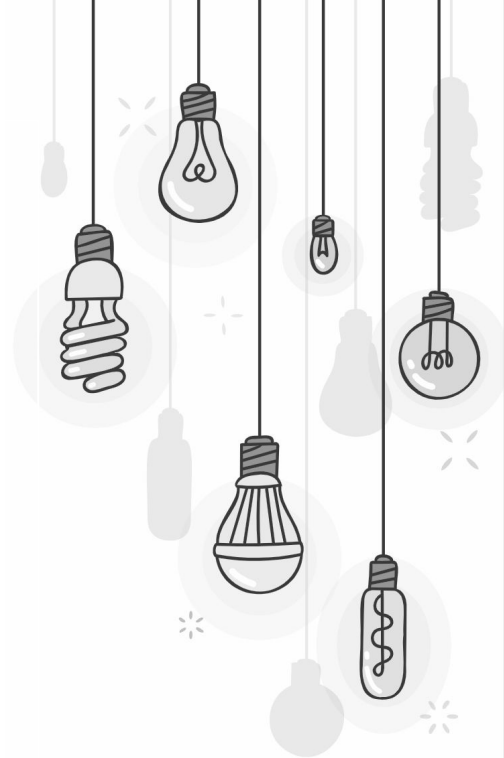
Figure 1: distribution of the implausible, plausible and neutral clarifications in CLAIRE.

Experiments

*How can we distinguish between
plausible and implausible
clarifications?*

Set-up

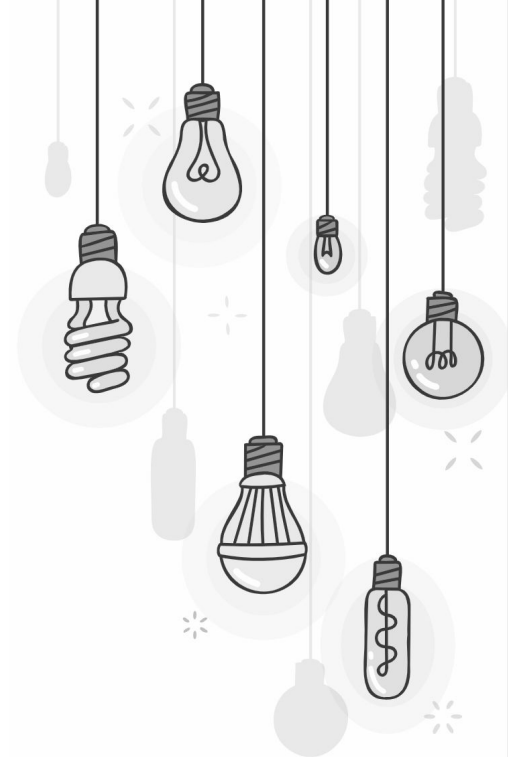
- * we approach the task as a **supervised classification** problem using our **train** ($N=19\,975$), **development** ($N=2500$) and **test set** ($N=2500$)
- * we use context before + clarified sentence + context after as **input**
- * we provided several **baselines** as a starting point
 - * **Naive Bayes**: BOW (unigrams)
 - * **BERT VANILLA**: BERT + linear classification layer
- * we **experiment** with several extensions of **BERT VANILLA** to mark the clarification



Results

Model	Accuracy
Naive Bayes	36.20 %
BERT VANILLA	44.51 %
BERT + RANKING FILLER MARKERS SENTENCE PAIR	48.53 % 51.39 % 50.68 %

Table 2: Accuracy score of the models on the development set.



Analysis

To what extent can clarifications reflect diverging interpretations of an implicit/underspecified element?

Set-up

- * We examine 159 instructions (*clarified sentence in context*) from the development set with at least **two plausible clarifications**
- * For each set of plausible clarifications, we examine if they were **conflicting** to one another or not

Conflicting clarifications are clarifications that refer to different entities/objects/aspects

- * In case of disagreement → discussion

Results

We found 116 instructions with **conflicting** clarifications

- * **Implicit references** (N=43)

- * **part-whole**: *refrigerate the oil for 1-2 weeks to infuse the oil/ingredients.*

- * **domain-related**: *Anger can only trouble your heart/life/soul.*

- * **Fused-heads** (N=27)

- * **Subgroup/group relationship**: *Most people/teenagers hate scary movies*

- * **Metonymy** (N=21)

- * **Different aspects**: *the absorption/the amount of sunlight.*

- * **Noun compounds** (N=25)

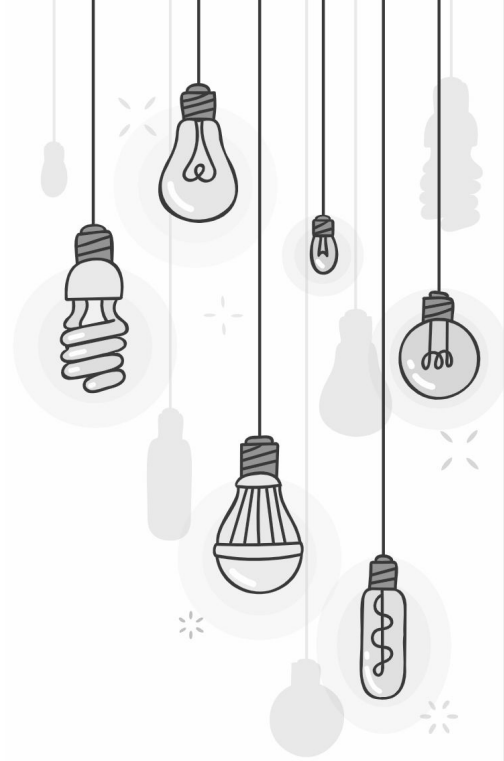
- * *the guitar/audio amps, summer/dance class, road/racing bike*

Conclusion

Conclusion

Summary

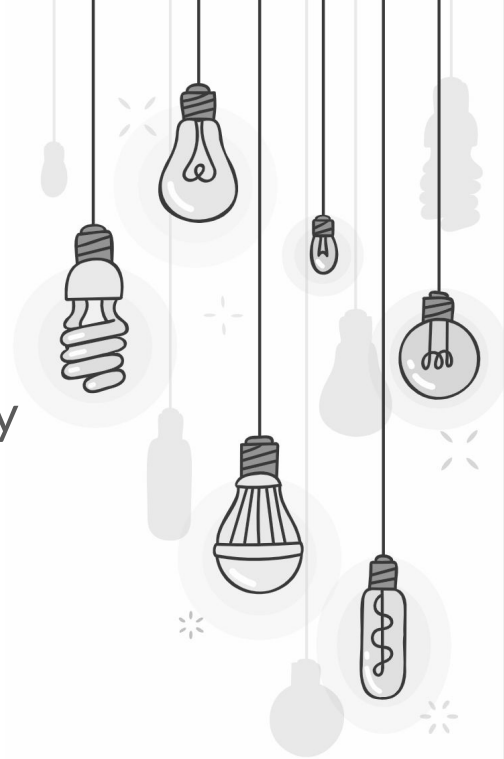
- * We released **CLAIRE**: a dataset of implausible and plausible clarifications for an instruction
- * We proposed the task of classifying a clarification as plausible, neutral or implausible for which we provided several **baselines**
- * We showed that some clarifications can be **plausible** but also **conflicting** to one another



Conclusion

Future steps

- * We plan to investigate the underlying reasons of why certain implicit and underspecified phrases have **multiple plausible clarifications**
- * We plan to find out how to model conflicting clarifications automatically



Thank you!

References

- * Antonio, T., & Roth, M. (2021). Resolving Implicit References in Instructional Texts. *Proceedings of the 2nd Workshop on Computational Approaches to Discourse (CODI)*, 58–71.
- * Antonio, T., Bhat, I., & Roth, M. (2020). wikiHowToImprove: A Resource and Analyses on Edits in Instructional Texts. *Proceedings of the 12th Language Resources and Evaluation Conference (LREC)*, 5721–5729.

