*[Work in progress]*
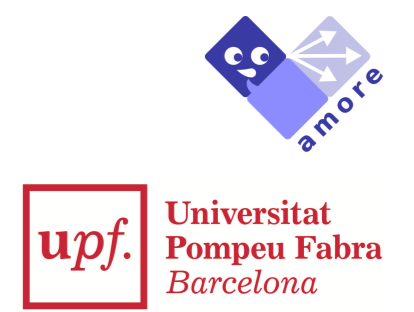*[hopefully]*

# Towards unsupervised language models for QUD prediction

## Matthijs Westera, Universitat Pompeu Fabra

Matthijs Westera, Universitat Pompeu Fabra

## Why QUD prediction?

- **Question Under Discussion (QUD)**: a set of relevant pieces of information that are jointly pursued [1,2].
- QUD is a very useful theoretical notion...
- ... but in practice QUD-based theories often require *explicit* questions to yield testable predictions.
- **Problem: QUDs are almost always *implicit*.**

## *Current approach*

### Data

- **LAMBADA** raw training data [3]:
  - 2.4K unpublished novels
  - 15M sentences (233M tokens)
  - around 1% (150K) ends with "?".
- Prefix sentences with tags *<say>*, *<ask>* based on punctuation (? vs ./!).

### Model *(for now...)*

Standard **neural** language model [4].
- Vocabulary: 50K×150 embeddings
- LSTM [5]: 2×500 units
- 30 epochs; backpropagate 130 tokens.

### Results *(for now...)*

For what it's worth (*some* hyperparameter optim.)
- Test **perplexity per word** overall: 140.25
  Questions only: 112.49
  *(i.e., model chooses right word as often as a 112-sided die.)*
- This isolated number doesn't mean much...
- Except perhaps that questions are more predictable than statements.

### Example output

**Prompt:**
   *"I carefully opened the box and looked inside. <ask>"*
**Generated:** (most likely 3-5 word questions from random sample):

| | |
|---|---|
| how did you know? | are you sure? |
| you don't know? | how did you know that? |
| you're not sure? | where are you? |
| you don't know what? | what's it? |
| what are you doing? | that's what? |
| what did you do? | I don't know? |
| where did you get? | is there anything else? |
| you want to go? | does it matter? |
| how did you know that? | is that what you think? |
| so , what was it? | can you see what? |

(more likely)

... many generic questions, only a few 'correct' ones.

## Related work

**Applications of QUD-based theories:**
- Exhaustivity / scalar implicatures [6]
- Negation [7]
- Intonation [2,8,9,10].
- Interpreting experimental results [11]
- Discourse coherence [2,10], cf. *rhetorical relations* [12]

**Question prediction** (among many)**:**
- Visual question prediction [13]
- LearningQ (from online forums) [14]

**QUD annotation:**
- Some exploratory work [15]

*[Suggestions welcome!]*

## *Some open issues:*

### Explicit questions $\xrightarrow{??}$ QUDs

- Are implicit and explicit questions sufficiently similar?
   *Suspicion: Yes, but explicit questions are more difficult to predict.*
- Explicit questions may explicate only part of a QUD.
- Not all 'questions' end with a "?".

### What sort of data to *train* on?

- Enough questions, sufficiently large, and sufficiently natural (so: ~~switchboard, wiki, news~~)
- Movie subtitles? Not self-contained...
- In fiction virtually all questions are in *reported speech*...

### What sort of data to *evaluate* on?

- QUD annotation? E.g., [15]. Costly and theory-laden.
- More natural (crowdsourcable) task: [work in progress]
   "which questions does this story evoke?"
- Secondary effects (e.g., intonation, exhaustivity).

## References

[1] Ginzburg (1996). Dynamics and the Semantics of Dialogue. *LL&C.*
[2] Roberts (1996). Information structure in discourse [...]. OSU WPL.
[3] Paperno et al. (2016). The LAMBADA dataset [...]. ACL.
[4] Pytorch examples. https://github.com/pytorch/examples .
[5] Hochreiter & Schmidhuber (1997). Long short-term memory. NC9.
[6] Westera (2016). An attention-based explanation [...]. *SuB21.*
[7] Kruszewski et al. (2016). [...] Conversational Negation [...]. CL42.
[8] Rooth (1992). A theory of focus interpretation. NLS1.
[9] Westera (2018). Rise-fall-rise [...]. In Gutzmann et al.
[10] Büring (2003). On D-trees, Beans, and Accents. L&P26.
[11] Westera & Brasoveanu (2014). Ignorance in context. *SALT24.*
[12] Hunter & Abrusan (2017). Rhetorical relations and QUDs. *LENLS.*
[13] Mostafazadeh et al. (2016). Generating nat. questions [...]. ACL54.
[14] Chen et al. (2018). LearningQ: A large-scale dataset [...]. AAAI.
[15] Riester et al. (2018). Annotation Guidelines [...]. In Adamou et al.