

Crowdfest Task 2

“Question Generation”

Rodrigo Agerri, Itziar Aldabe, Marie Bexte,
Andrea Horbach, **Montse Maritxalar**, Nadia Okinina

enetCollect – European Network for Combining Language Learning with
Crowdsourcing Techniques

15th March, Lisbon



enetCollect

COST Action CA16105

<http://enetcollect.eurac.edu/>



COST is supported by the EU
Framework Programme
Horizon 2020

Aim

to create an interdisciplinary
community around the
generation of pedagogical
questions that should be
relevant in the task of reading
comprehension.

Overview

- Objectives
- Evaluation of Automatically Generated Questions
- Annotation Guidelines
 - Preliminary Data Analysis
 - Evaluating the criteria
- Evaluation of Manually Generated Questions
- Lessons learnt



Objectives

- Evaluation of questions via crowdsourcing
- Hackathon:
work on the improvement of the QG task



Evaluation of Automatically Generated Questions

- Biology texts (grade 7/8): Cell theory
- English
- Rule-based approach and deep learning approach
- Evaluation platform: pybossa
- 5 measures: Fluency, ambiguity, answerability, pedagogical relevance, comprehension type



Evaluation of Automatically Generated Questions

Examples

- When did Rudolf Virchow develop the hypothesis that cells only come from other cells?
- What is a cell?
- What is a human to a potato?
- What is so?
- What do many structures in the cell have?
- Who created the theory of the theory?



Annotation Guidelines

PyBossa · Project: Cell Theory · Contribute - PyBossa by Scifabric - Mozilla Firefox

PyBossa · Project: Cell Theory

pybossa.ixax.eu/project/CT_QG/task/565

pybossa Community Projects Create About montse.maritxalar

Cell Theory: Contribute

Text

A cell is the smallest structural and functional unit of an organism. Some organisms, like bacteria, consist of only one cell. Big organisms, like humans, consist of trillions of cells. Compare a human to a banana. On the outside, they look very different, but if you look close enough you'll see that their cells are actually very similar.

Most cells are so small that you cannot see them without the help of a microscope. It was not until 1665 that English scientist Robert Hooke invented a basic light microscope and observed cells for the first time, by looking at a piece of cork. You may use light microscopes in the classroom. You can use a light microscope to see cells (Figure below). But many structures in the cell are too small to see with a light microscope. So, what do you do if you want to see the tiny structures inside of cells?

In 1858, after using microscopes much better than Hooke's first microscope, Rudolf Virchow developed the hypothesis that cells only come from other cells. For example, bacteria, which are single-celled organisms, divide in half (after they grow some) to make new bacteria. In the same way, your body makes new cells by dividing the cells you already have. In all cases, cells only come from cells that have existed before. This idea led to the development of one of the most important theories in biology, the cell theory.

As with other scientific theories, many hundreds, if not thousands, of experiments support the cell theory. Since Virchow created the theory, no evidence has ever been identified to contradict it.

Question

what do many structures in the cell have ?

Annotation Guidelines

Fluency

- The question is grammatically correct and idiomatic/natural
- The question is grammatically correct but does not read as fluently as we would like.
- There are some grammatical errors in the question
- The question is grammatically unacceptable.

Ambiguity

- The question itself is completely unambiguous
- The question could provide more information
- The question is totally ambiguous when asked out of the blue

Answerability

- The question can be answered with the information from the source text.
- The question can be answered with the information from the source text plus some world knowledge*.
- The question can be answered using some world knowledge*, and without using information from the text
- The question can not be answered with the information from the source text, neither using some world knowledge*.

Pedagogical relevance

- The question is compulsory to be sure that the reader has understood the text.
- The question is relevant to evaluate the reading comprehension level but it is not compulsory.
- The question is related to the topic of the text but it does not measure its reading comprehension level.
- The question is not related to the topic of the text.

Comprehension type

- The question can be answered directly and explicitly from the text (facts, dates, times, locations, entities...)
- It is necessary to combine information from various parts of the text in order to answer the question.
- It is necessary to combine information from the literal understanding of the text with reader's knowledge and intuitions in order to answer the question.
- The question is about what might happen next or after a story ends.
- The question asks about giving a global or comprehensive judgement about some aspect of the text.
- The question is about the reader's feelings for the text and the subject.

Preliminary Data Analysis

- How do human-annotated quality scores correlate with properties of the generated data?
- **Ultimate Goal:** automatically rank questions or exclude bad questions
- **Properties considered:**
 - Question length
 - Overlap of the question with the source sentence (answer)
 - n-gram probability of the question
- **Number of evaluated questions: 28**

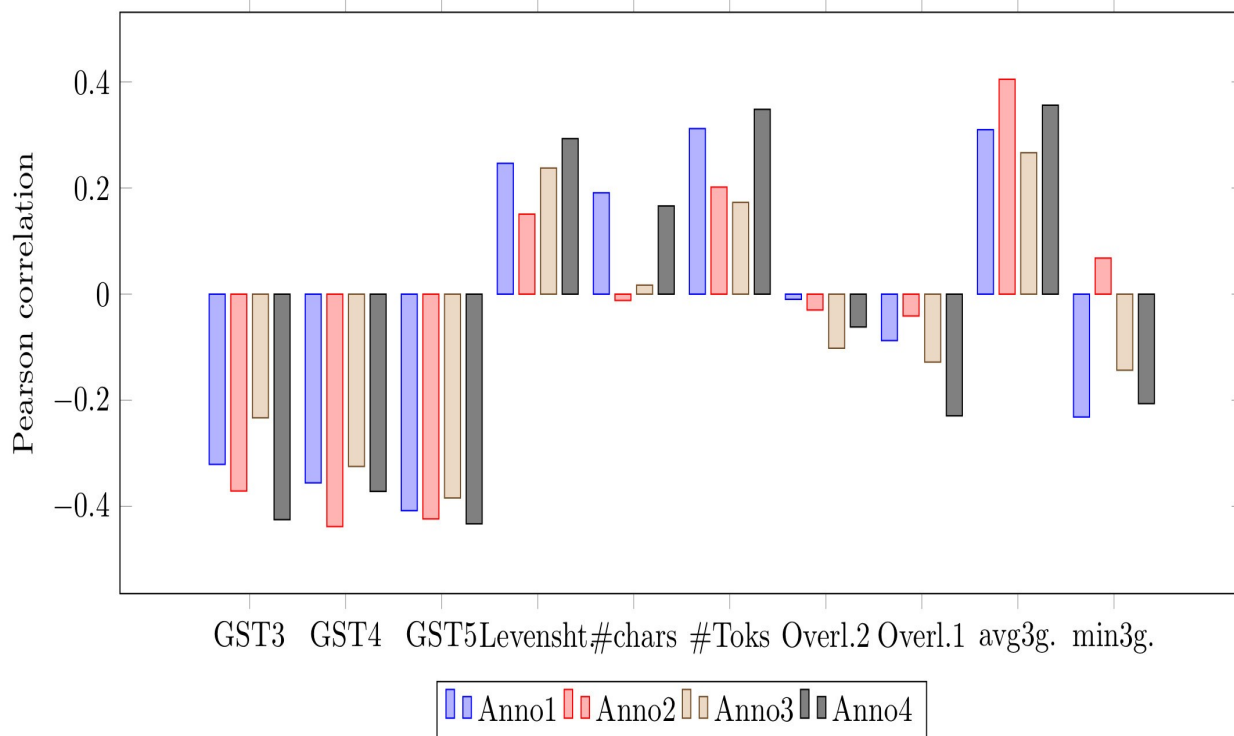


Data Analysis for Fluency

- **GST (Greedy String Tiling):** the higher the similarity the lower fluency
- **Levenshtein distance:** opposite to GST
- **#Toks:** in the question with no high correlation
- **Overl1:** % #toks of the question in the answer
- **Overl2:** % #toks of the answer in the question
- **Avg3g:** 3-gram (question) probability under a generic Web1T English language model
- **Min3g:** higher probability more fluent

Lower fluency means better

Pearson Correlation for FLUENCY
Individually for each of the 4 annotators



Preliminary Data Analysis

Fluency

Annotator1

GST3 feature

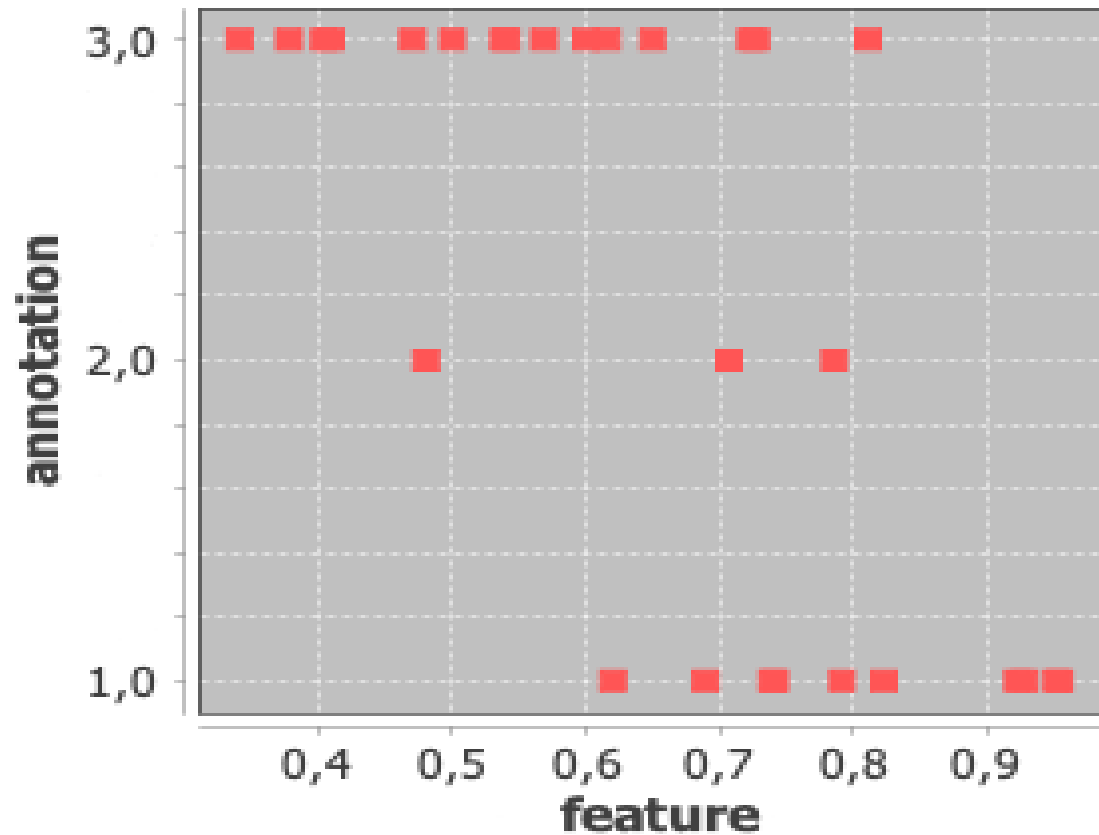
REMEMBER:

The higher the similarity
the lower fluency

and

lower fluency means
better

Scatterplot



Annotation Guidelines: Evaluating the criteria

- Low inter-annotator agreement
- Depending on the question, some evaluation criteria were not applicable
- New evaluation guidelines
- New pybossa task presenter to show some criteria based on previous answers



New Annotation Guidelines

- Same questions
- External evaluation of an expert
- New guidelines
 - Remove idiomatic
 - Rephrasing of the question



New Annotation Guidelines

PyBossa · Project: crowdfest_guidelines · Contribute - PyBossa by Scifabric - Mozilla Firefox

PyBossa · Project: crowd X

pybossa.ixax.eu/project/CT_QG_guidelines/task/804

90%

Search

montse.maritxalar

A ANALYZE THE NEXT QUESTION. **X**

Text

A cell is the smallest structural and functional unit of an organism. Some organisms, like bacteria, consist of only one cell. Big organisms, like humans, consist of trillions of cells. Compare a human to a banana. On the outside, they look very different, but if you look close enough you'll see that their cells are actually very similar.

Most cells are so small that you cannot see them without the help of a microscope. It was not until 1665 that English scientist Robert Hooke invented a basic light microscope and observed cells for the first time, by looking at a piece of cork. You may use light microscopes in the classroom. You can use a light microscope to see cells (Figure below). But many structures in the cell are too small to see with a light microscope. So, what do you do if you want to see the tiny structures inside of cells?

In 1858, after using microscopes much better than Hooke's first microscope, Rudolf Virchow developed the hypothesis that cells only come from other cells. For example, bacteria, which are single-celled organisms, divide in half (after they grow some) to make new bacteria. In the same way, your body makes new cells by dividing the cells you already have. In all cases, cells only come from cells that have existed before. This idea led to the development of one of the most important theories in biology, the cell theory.

As with other scientific theories, many hundreds, if not thousands, of experiments support the cell theory. Since Virchow created the theory, no evidence has ever been identified to contradict it.

Question

what can be used to describe cells in a light bulb ?

Answer the next criteria

Is it understandable? Yes No

New Annotation Guidelines

PyBossa · Project: crowdfest_guidelines · Contribute - PyBossa by Scifabric - Mozilla Firefox

PyBossa · Project: crowd X

pybossa.ixa.eus/project/CT_QG_guidelines/task/804

pybossa Community Projects Create About montse.maritxalar

PLEASE SUBMIT THE NEXT CRITERIA.

Question

what can be used to describe cells in a light bulb ?

Answer the next criteria

Is it understandable? Yes No

Submit

Is it domain-related? Yes No

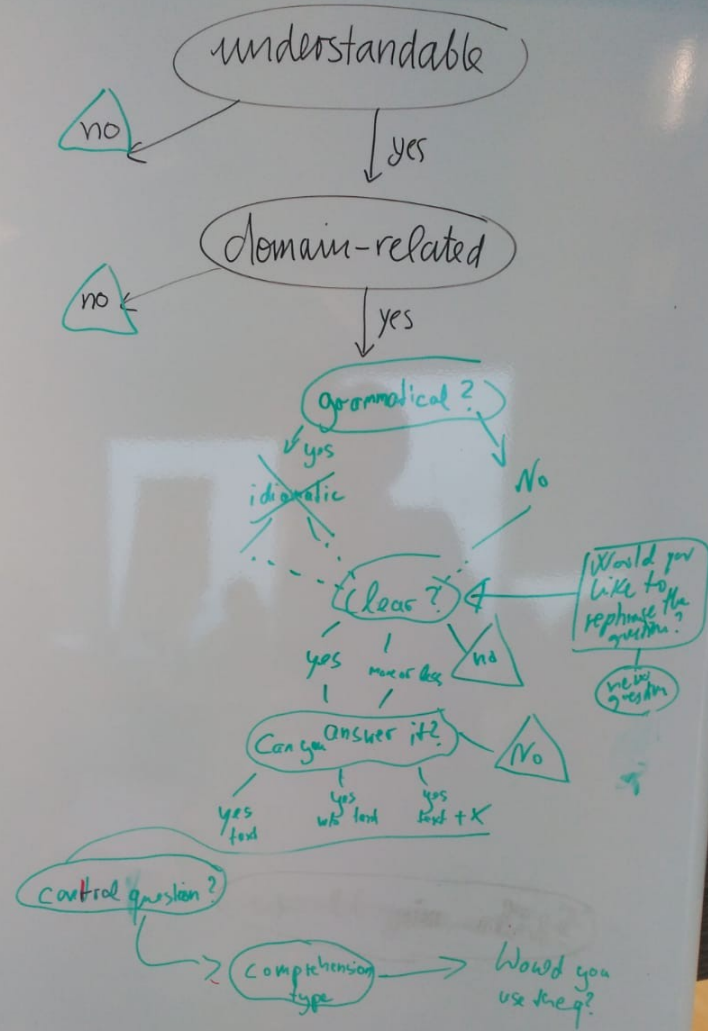
Submit

Is it grammatical? Yes No

Is the question clear? Yes More and less No

Submit

New Annotation Guidelines



Evaluation of Manually Generated Questions

- Created by a group member

Criteria:

- Not be too boring
 - Paraphrase
 - Ask questions only about the content of the text
 - Some questions require short and precise answers
 - Other questions allow longer and more personalised answers
 - Cover the content of the whole text
 - Use different syntactic constructions
- Evaluated by the rest of the group
 - Inter-annotator agreement (to analyse)



Evaluation of Manually Generated Questions

Examples

- How do our bodies produce new cells?
- What do a human and a banana have in common?



Lessons learnt

- Annotation/Evaluation is challenging
- Linguistic measures are good as a first step to measure the quality of the questions
- We need measures to evaluate the quality of the questions from the pedagogical point of view
- Need more feedback from educational experts



References and Outcomes

OUTCOMES

- Automatically generated Question Dataset
- Pybossa platform
- Implemented Guidelines (improving)

MAIN REFERENCES

- Du, X., Shao, J., & Cardie, C. (2017). Learning to ask: Neural question generation for reading comprehension. arXiv preprint arXiv:1705.00106.
- M. Heilman and N. A. Smith. 2010. Good Question! Statistical Ranking for Question Generation. In Proc. of NAACL/HLT.



You can check some automatically generated questions using

The first version of the annotation guidelines at:

http://pybossa.ixxa.eus/project/CT_QG/task/556

New guidelines at:

http://pybossa.ixxa.eus/project/CT_QG_guidelines/task/812

Thank You

Montse Maritxalar
montse.maritxalar@ehu.eus