

Factual Question Generation for the Portuguese Language

B. Leite (bernardo.leite@fe.up.pt), H. L. Cardoso, L. P. Reis, C. Soares
Artificial Intelligence and Computer Science Lab
University of Porto, Portugal

INISTA 2020 - International Conference on
INnovations in Intelligent SysTems and Applications
August 24-26, 2020 Novi Sad, Serbia

Agenda

- **Introduction**
- Problem
- Goals
- Literature Review
- Methodology
 - Syntax-based
 - Semantic-based
 - Dependency-based
- Evaluation
- Conclusions and Future Work

Introduction



- **Artificial Intelligence (AI) is changing the world**, either the way we live, work, or learn;
- The area of education is also changing: **There is a growth in the use of AI in teaching and learning** from primary to higher education;
- Educational tools provided by AI demonstrate their potential **to improve the quality of teaching** and traditional learning methods.

Problem



The act of questioning is one of the most used traditional methods by teachers to assess students' knowledge or understanding. However...

- It proves to be a **very time-consuming task**.
- Taking into account the difficulty in creating new and rich questions, these are often **repeated**, already **used in previous contexts**.

Question Generation (QG) has shown to be one of the most important applications for optimizing this process.

Goals



- Apply **Artificial Intelligence** (AI) and **Natural Language Processing** (NLP) techniques to **generate factual questions** from Portuguese educational texts;
- **Adapt and improve approaches previously proposed in the literature** (syntax-based, semantic-based and dependency-based);
 - For a specific language - Portuguese.
- **Give a sense that the proposed methods can be applied to other contexts**, regarding other languages.

Agenda

- Introduction
- Problem
- Goals
- **Literature Review**
- Methodology
 - Syntax-based
 - Semantic-based
 - Dependency-based
- Evaluation
- Conclusions and Future Work

Literature Review - General Overview

Types of questions:

- multiple-choice;
- factual (How? Who? Where? When?);
- gap-fill;
- true/false.

Domains of questions:

- Biology, medical sciences, geometry, physics, maths, computer science;
- Language learning.

Language of questions:

- English, Basque, French, Chinese, Russian;
- Portuguese

- *[Pirovani et al., 2016], [Correia et al., 2012], [Diéguez et al., 2011], [Curto, 2010]*

Literature review - Different Approaches

- **Template-based** by using pre-defined templates of the questions;
 - [Awad et al., 2014], [Le et al., 2015]
- **Syntactic Analysis** by manipulating the syntactic structure of the sentence;
 - [Majumder et al., 2015], [Danon et al., 2017]
- **Semantic Analysis** which focus on semantic parse (using Semantic Role Labeling);
 - [Araki et al., 2016], [Flor et al., 2018]
- **Dependency Analysis** which connects words in a sentence using their functional relations;
 - [Mazidi et al., 2016]
- **Machine Learning** techniques using Neural Networks (*Seq2Seq* approaches).
 - [Chen et al., 2018], [Lu et al., 2019]

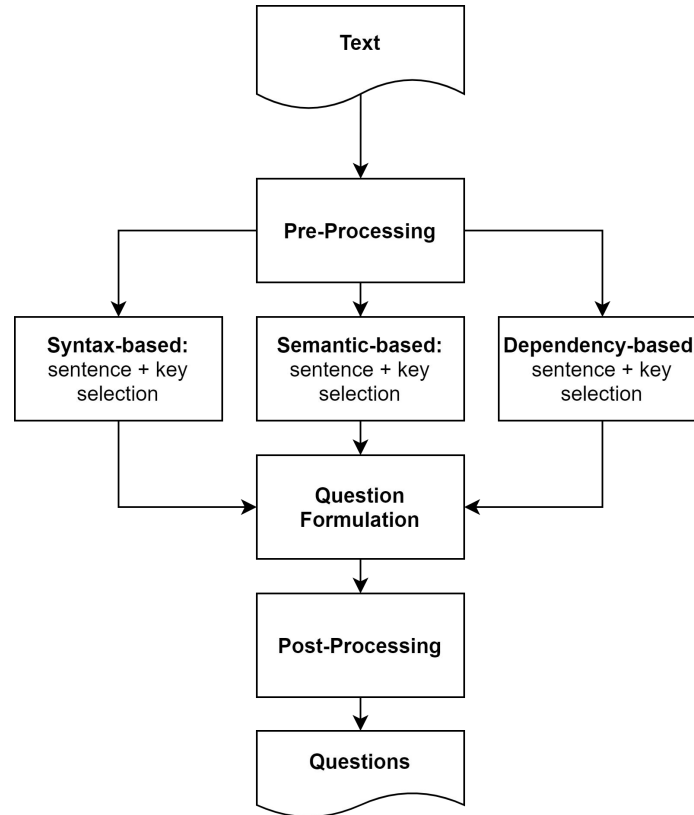
Literature review - *How to evaluate the quality of the computer-generated questions?*

- There is **no standard form** in order to evaluate the results for question generation;
- Most often what can be found in the literature is **manual evaluation** of the computer-generated questions;
- **Various metrics** have been employed by researchers depending on the application and domain. [Kurdi et al., 2019]
 - Relevance, acceptability, answerability, grammatical correctness, semantic ambiguity, fluency, number of errors.

Agenda

- Introduction
- Problem
- Goals
- Literature Review
- **Methodology**
 - Syntax-based
 - Semantic-based
 - Dependency-based
- Evaluation
- Conclusions and Future Work

Methodology: General Overview



Methodology: Tools/Technologies

- **NLTK** to perform
 - Tokenization: The process of breaking a stream of text up into words, phrases, symbols, or other meaningful elements called tokens.
- **spaCy** to perform
 - Part-of-Speech Tagging (PoS): The process of marking up a word in a text as corresponding to a particular part of speech (nouns, verbs, adjectives, adverbs, etc).
 - Named Entity Recognition (NER): Seeks to locate and classify named entities.
 - Dependency Parsing: To assign dependency labels within the words that compose our sentences.
- **nlpnet** to perform
 - Semantic Role Labeling: Identifies the semantic roles of each argument and modifier in a sentence (time, manners, locations, etc).

(1) Syntax-based Approach - Using *PoS* and *NER*

Procedures	Results
Select sentence(s) (with 1 or more entities)	Francisco Pizarro descobriu o Império Inca na América do Sul . (<i>Francisco Pizarro discovered the Inca Empire in South America.</i>)
Produce morphological sequence (from PoS):	[NOUN, NOUN, VERB, DET, NOUN, NOUN, PREP, NOUN, PREP, NOUN, PUNCT]
Combine entities (from NER) with morphological sequence	[PER , PER , VERB, DET, NOUN, NOUN, PREP, LOC , LOC , LOC , PUNCT]
Search patterns (using pre-established rules)	[PER][VERB AUX].*?[PUNCT] [PER][VERB AUX].*?[LOC][PUNCT]
Formulate questions (applying the necessary transformations)	Quem descobriu o Império Inca na América do Sul? (<i>Who discovered the Inca Empire in South America?</i>) Onde é que Francisco Pizarro descobriu o Império Inca? (<i>Where did Francisco Pizarro discover the Inca Empire?</i>)
These questions can start with: Who (Key: PER), Where (Key: LOC), Which organization (Key: ORG)	

(2) Semantic-based Approach - Using Semantic Role Labeling

Procedures	Results
Select sentence(s) (only if the intended semantic roles are identified)	No dia seguinte, o rei estava muito fraco e cansadíssimo . (<i>The next day, the king was very weak and tired.</i>)
Semantic Roles	<i>A0: the king</i> <i>V: was</i> <i>AM-MNR: very weak and tired</i>
Search patterns (using pre-established rules)	[A0 V A1][V][AM-MNR]
Apply transformation	“Como (<i>How</i>)” + [V] + [A0] + “?”
Formulate question	Como estava o rei? (<i>How was the king?</i>)
These questions can start with: Who (Key: A0 V A1), Where (Key: AM-LOC), When (Key: AM-TIME) and How (Key: AM-MNR)	

(3) Dependency-based Approach

Procedures	Results
Select sentence(s) (only if the desired dependencies are identified)	Nenhum peixe-voador rasgava a superfície , e não havia redemoinhos de peixe miúdo. (No flying fish tore the surface , and there were no swirls of small fish.)
Dependencies:	(flying fish, nsubj), (tore, root), (the, determiner), (surface, obj)
Produce sequence from the dependencies	[det, nsubj, root, det, obj, punct, cc, advmod, conj, obj, case, nmod, amod, punct]
Search patterns (using pre-established rules)	[det, nsubj, root, det, obj]
Apply transformation	"O que é que (What)" + [det] + [nsubj] + [root] + "?"
Formulate question	O que é que nenhum peixe-voador rasgava? (What did no flying fish tear?)
Using Direct Object we can generate questions starting with O que é que (What) Using Indirect Object we can generate questions starting with A quem é que (To whom)	

Dependency-based Approach - More Examples

- **Sentence:**

*Os gatos **são maus**, alguns foram apanhados em flagrante almoçando andorinhas, havia alguma verdade nisso.*

*(Cats **are bad**, some were caught in blatant having lunch swallows, there was some truth to that.)*

subject + verb *to be* + adjective

This is an indicator that we can have a **characterization** in our sentence.

- **Question:**

Como caracteriza os gatos?

*(**How do you characterize** cats?)*

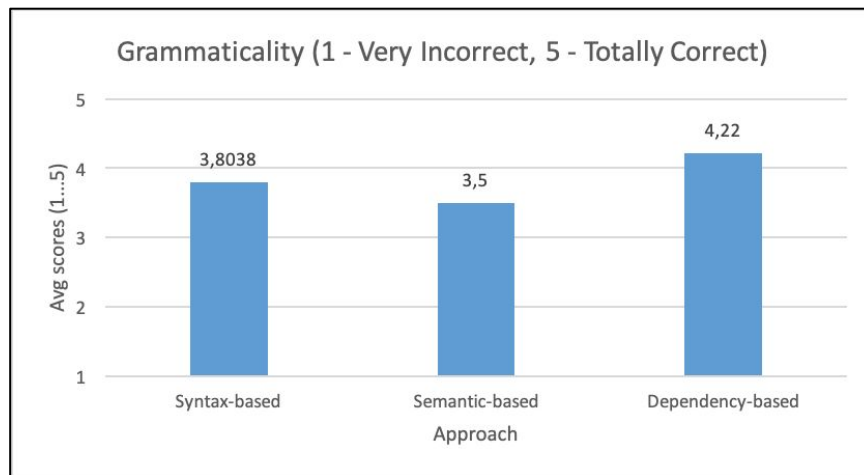
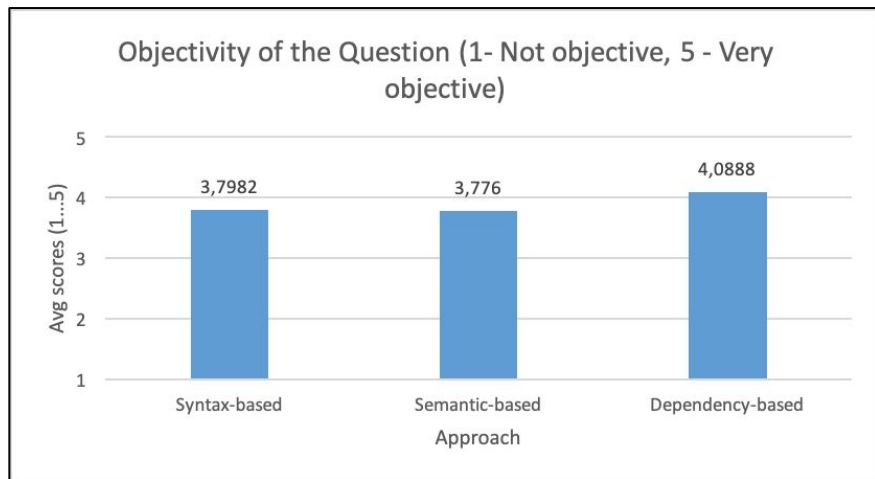
Agenda

- Introduction
- Problem
- Goals
- Literature Review
- Methodology
 - Syntax-based
 - Semantic-based
 - Dependency-based
- **Evaluation**
- Conclusions and Future Work

Evaluation - Teacher Survey

- **Survey** (10 Portuguese Teachers)
 - **15** generated questions from **4** Portuguese books (belonging to National Reading Plan);
 - Syntax-based approach: **5** questions
 - Semantic-based approach: **5** questions
 - Dependency-based approach: **5** questions
- **Metrics** (5 metrics per question)
 - **Objectivity** of the Question - Do you consider the question objective?
(1 - Nothing objective, 5 - Very objective)
 - **Grammatically** - Do you consider the question to be grammatically correct?
(1 - Very Incorrect, 5 - Totally Correct)
 - **Question Extension** - Do you consider the length of the question adequate?
(1 - Not long, 5 - Too long)
 - **Answerability** - How many answers do you think this question might have?
(No answer, One, Two or more)
 - Please **rewrite the question** to a correct form.
(if needed)

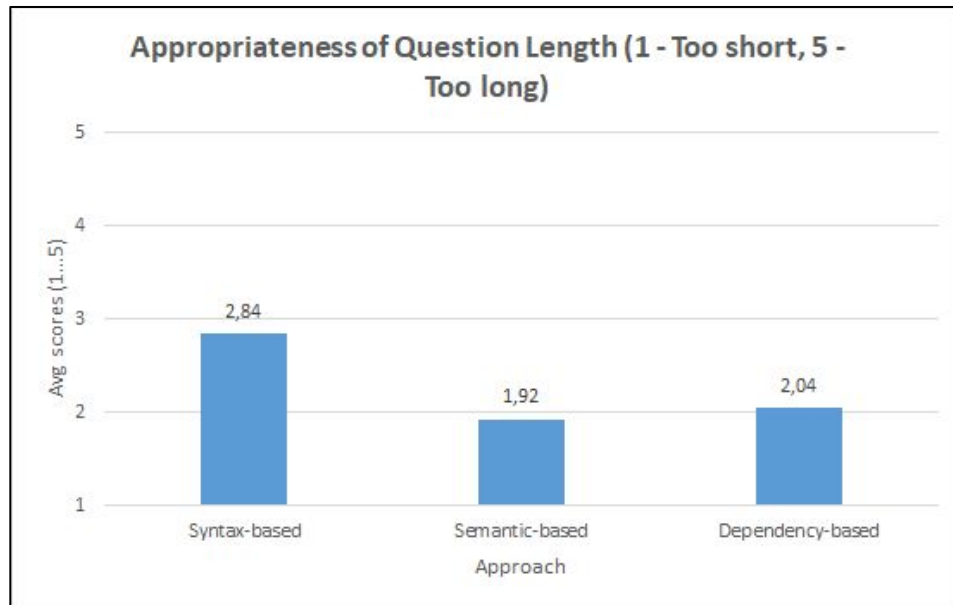
Results: Objectivity and Grammaticality



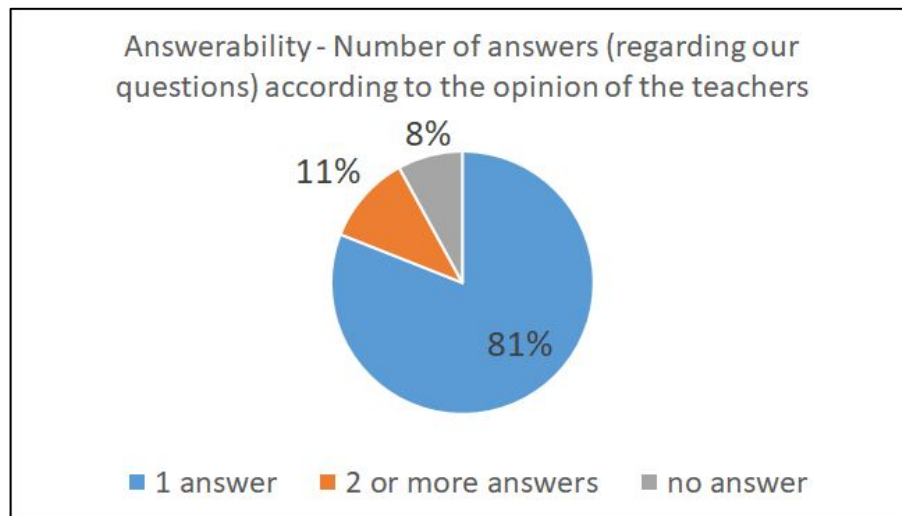
- Overall the Teachers found the **questions objective**;
- **Best results go for dependency-based** approach, because this type of questions is quite incisive by using objective terms like **What**.
- The main errors that can arise are related with the **conjugation of verb tenses**;
- Lowest score for the semantic-based approach
 - Questionable term is not always the most appropriate.

Results: Appropriateness of Question Length

- We consider that the appropriateness of question length has an **optimal value of 3**;
- The size of the question **is adequate most of the time**;
- **Semantic-based and dependency-based** approaches have a **lower value** for this metric
 - Generated questions are typically shorter.



Results: Answerability - Number of answers



There are different opinions on the correct number of answers

- Number of facts (or entities) as possible answers;
- Multiple adjectives may be available for questions that require characterizations.

Results: Rewriting the question to a correct a form (suggestions by teachers)

- In most cases, teachers suggested to
 - **simplify** the questions;
 - **add auxiliary verbs**;
 - **slightly change the order of some words**.
- Regarding **Avg. Levenshtein**, the **similarity** between generated and suggested questions is between **70 to 80 percent**.

(1) Conclusions

- **Syntax-based**

- It provides a **satisfactory understanding of the sentence**: NER and PoS tags;
- When combined with NER, allows the creation of **very specific questions**.

- **Semantic-based**

- It **captures the semantic information** from the sentences;
- It has the advantage of raising questions about **relevant information** such as manners, times and locations.

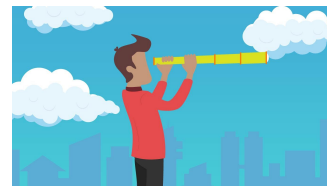
- **Dependency-based**

- It takes advantage of certain **dependency relations** (i.e., direct and indirect object);
- It allows to identify **characterizations** of things and people in declarative sentences.

(2) Conclusions

- Question Generation proves to be a **promising field** to improve **education process**;
- A combination between syntax, semantic and dependency-based approaches **provide various and rich questions**;
 - Also, the proposed methods and their procedures can be applied to other languages.
- We were able to **fill gaps from the literature** namely...
 - the little study in the generation of questions for Portuguese.
 - the lack of patterns/rules to identify new questionable facts.
- Our findings outline a promising future for QG, **not only in Portuguese but for other languages that contain little study.**

Future Work



- Generate questions from **multi-line facts**;
- Control **the difficulty of the questions**;
- **Personalize questions** by student needs;
- Use of **standard evaluation techniques**;
- Try new approaches for AQG such as **Machine Learning techniques** but...
 - always ensuring the **educational relevance** of the questions.

Some progress has already been made in my master's thesis

Bernardo Leite, Henrique Lopes Cardoso and Luís Paulo Reis. Automatic Question Generation for the Portuguese Language. Master's thesis, Faculdade de Engenharia da Universidade do Porto, 2020.

Thank you!

B. Leite (bernardo.leite@fe.up.pt), **H. L. Cardoso, L. P. Reis, C. Soares**
Artificial Intelligence and Computer Science Lab
University of Porto, Portugal

INISTA 2020 - International Conference on
INnovations in Intelligent SysTems and Applications
August 24-26, 2020 Novi Sad, Serbia