

# Investigating Methods for Automatic Question Generation from Lecture Transcripts

## Overview

- Question generation from lecture transcripts is a useful tool to assist lecturers anticipate what questions may be asked and to allow students to learn through inquiry-based learning.

## Problem Statement

- Determining whether high-caliber questions can be generated using a small, high-quality dataset, either through a neural network or template-based system.

## Objectives

### Semantic & Template Systems:

- Build a Semantic Rule-Based System (Baseline).
  - Rules are manually created.
- Build an Automatic Template-Based System.
  - Automatic template extraction from input questions
- Aims to automate the tedious process of template creation while improving coverage and maintaining quality vs manual semantic system.

### Neural System:

- Fine-tune t5-base and docT5query on the dataset of lecture transcripts and student questions (Baseline).
- Determine if similar performance can be achieved on a smaller dataset.
- Determine if exposing the models to further question-context pairs can improve the generated questions.

## Semantic & Template Systems

### Transcript Pre-Processing

Contractions are expanded and the transcript is segmented into sentences.

### Match to Rules/Templates

Sentences undergo semantic role labelling (SRL). Identified tags are matched to the Rules/Templates.

### Output Questions

Sentence: The boy kicks the ball.  
SRL Tags: <arg0> <V> <arg1>.  
Template: Who <V> <arg1>?  
Question: Who kicks the ball?

## Neural System

- Finetune the language model (t5-base) and the pre-trained model (docT5query) on the following datasets:

### IR-Small

Diverse set of 17 information retrieval queries and documents datasets.

### LearningQ

Contains questions and context pairs from an educational platform.

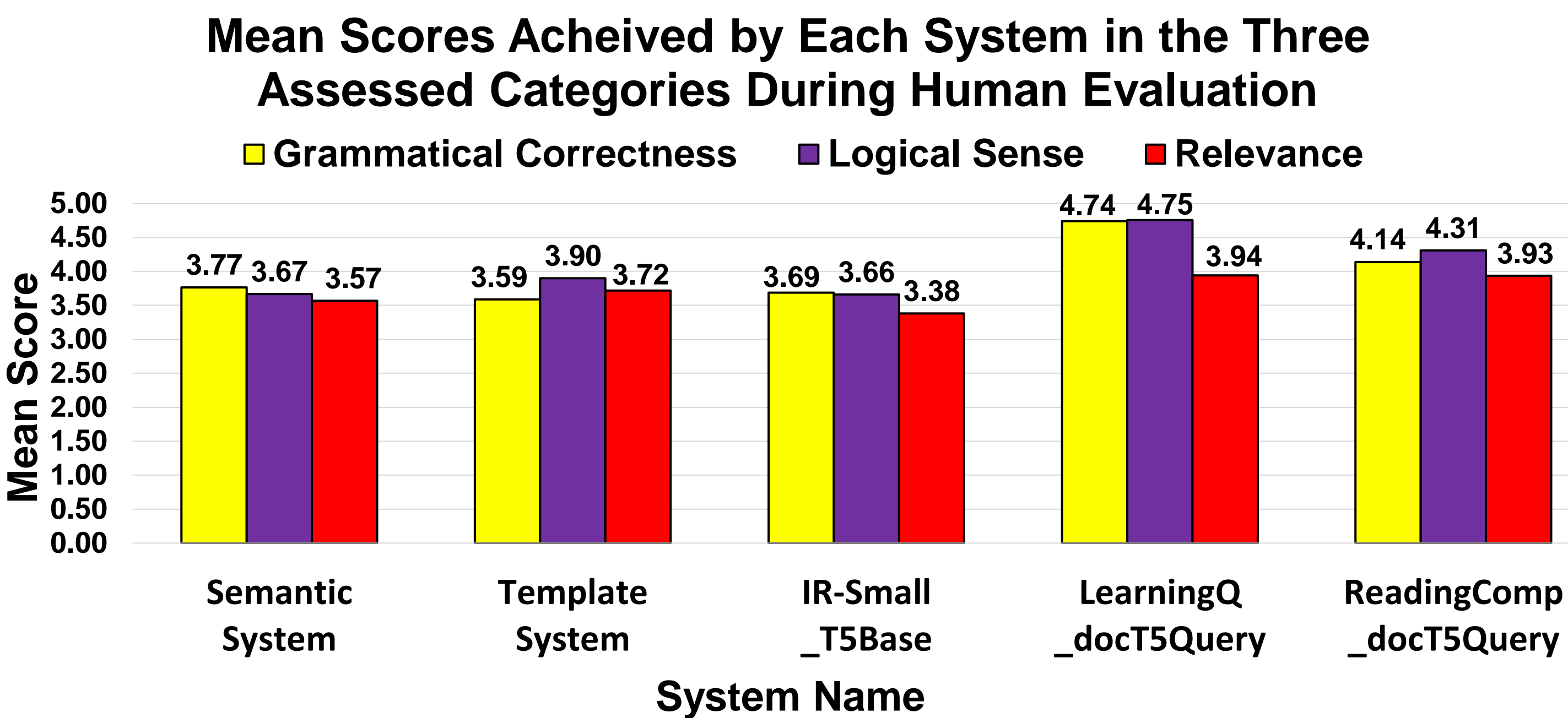
### ReadingComp

A collection of 5 question-answer datasets and their contexts.

## Results

### Assessment Procedure:

- Question quality was assessed using an online survey.
- 15 participants received 10 contexts, each accompanied by 5 questions generated by the different systems.
- Questions were scored (1-5) in the following 3 categories:
  - Grammatical Correctness
  - Logical Sense
  - Relevance

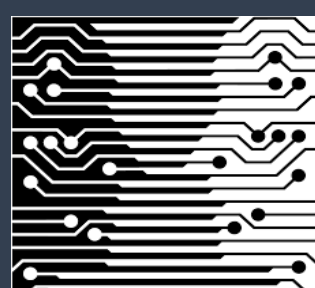


## Conclusions

- Template-based system matched the output quality of the semantic system with improved coverage.
- IR-Small\_T5Base performed worse than the docT5Query baseline.
- Training on LearningQ enhanced performance relative to the baseline.
- Both approaches generated high-quality questions with the LearningQ neural model being favoured by survey participants.

## Future Work

- Template:** Introduce a categorisation system to enable better logically linking of templates and sentences.
- Neural:** Experiment with word embeddings and linguistic features.



Supervisor  
Dr. Zola Mahlaza

University of Cape Town  
Department of Computer Science  
Telephone: +27 (0)21 650 2663  
E-Mail: dept@cs.uct.ac.za

Project Team  
Liam Talberg  
Adam Vere

