Question Answering and Lifelong Learning

Programme: Intelligent Systems

Guillermo Echegoyen

Advised by:

Anselmo Peñas

Alvaro Rodrigo

Table of contents

Who am I

Who am I

Background

- Software Engineering
- AI Master

Where to find me

- Resumé
- Google Scholar

Currently

- International project LIHLITH-KIQA (PI: Anselmo Peñas). Question Answering that evolve over time.
- First year PhD in Lifelong Learning Question Answering.

Introduction

Frame: LIHLITH project

LIHLITH-KIQA, IP: Anselmo Peñas International coordinator: Eneko Agirre

European funded project. PI: by Eneko Agirre.

Learning to Interact with Humans by Lifelong Interaction with Humans

- Develop Question Answering and Dialogue Systems
- Enable them to evolve over time.

Question Answering

Systems that automatically answer questions posed by humans in a natural language

- Open world assumption
- Closed world assumption

Applications:

- Question Answering Systems are everywhere, from personal assistants to chatbots and IT ticket management.

Lifelong Learning

Adapt previous knowledge to solve new, unseen problems.

Can be combined with many algorithms and different forms of learning:

- Neural networks
- Topic modeling
- Information Extraction
- Reinforcement Learning

Research Plan

Objectives

- Develop systems capable of detecting knowledge gaps
- Develop strategies to fill these gaps
- Enable QA systems to evolve over time

Research Questions

- How to detect unanswerable questions given the available knowledge
- How to incorporate new knowledge with previous one
- How to asses systems effectiveness

Research Plan

1rst Year

- Study state-of-the-art
- Gather evaluation collections
- Propose QA-LL system
- Propose method to develop evaluation collections

2nd Year

- Develop collections for LL QA
- Evaluation and analysis of results
- Improvements from results
- Detect new challanges

Research Plan

3rd Year

- Development of QA-LL System
- Evaluation and analysis of results
- Final draft with implemented system

4rd Year

- Evaluation and final touches
- Write thesis

Outcomes

Publications (older to newer):

- QA-LL task proposal: [3]
- First evaluation collection: [4]
- Survey on dialogue systems: [1]
- Cross-lingual models: [2]
- Lifelong Learning Transformers (work in progess)

Question Answering-Lifelong Learning task proposal

Given a KG, and an external resource such as a text collection in the same domain:

1. Decide whether a user utterance in Natural Language can be mapped or not into the KG and, if this is the case
2. Determine through which ones of the following ways the system must find a strategy to enrich the knowledge graph

Outputs

Shared task

First evaluation collection:

- Extract information from Wikimedia Cookbook.
- Organize and structure information as Knowledge Graph.
- Manually create both answerable and unanswerable questions from the KG.
- Manually annotate questions and answers.

Outputs

- Cooking Domain Knowledge graph
- Cooking Domain Question Answering dataset

Survey on Dialogue Systems

Study the state-of-the-art in:

- Task Oriented Dialogue System
- Conversational Dialogue Systems
- Question Answering Dialogue Systems
- Evaluation Datasets and Challenges

Cross Lingual Models

Goal: Train a model in a language and transfer learning to another one (and another task).

Dataset	BERT	MultiBERT	Random	Longest
RACE Mid	0.5265	**0.6114**	0.2500	0.3078
RACE High	0.4774	**0.5031**	0.2500	0.3059
RACE All	0.4917	**0.5347**	0.2500	0.3059
EE English	0.4921	**0.4974**	0.2500	0.2304
EE Spanish	0.3665	**0.4503**	0.2500	0.2932
EE Italian	0.2880	**0.4293**	0.2500	0.2775
EE French	0.3037	**0.4346**	0.2500	0.2565
EE Russian	0.2618	**0.3403**	0.2500	0.2723
EE German**	0.3708	**0.4494**	0.2500	0.2584

Lifelong Learning Transformers

Goal: Reuse knowledge from previous tasks to boost performance of a new one while preserving effectiveness.

ToDo

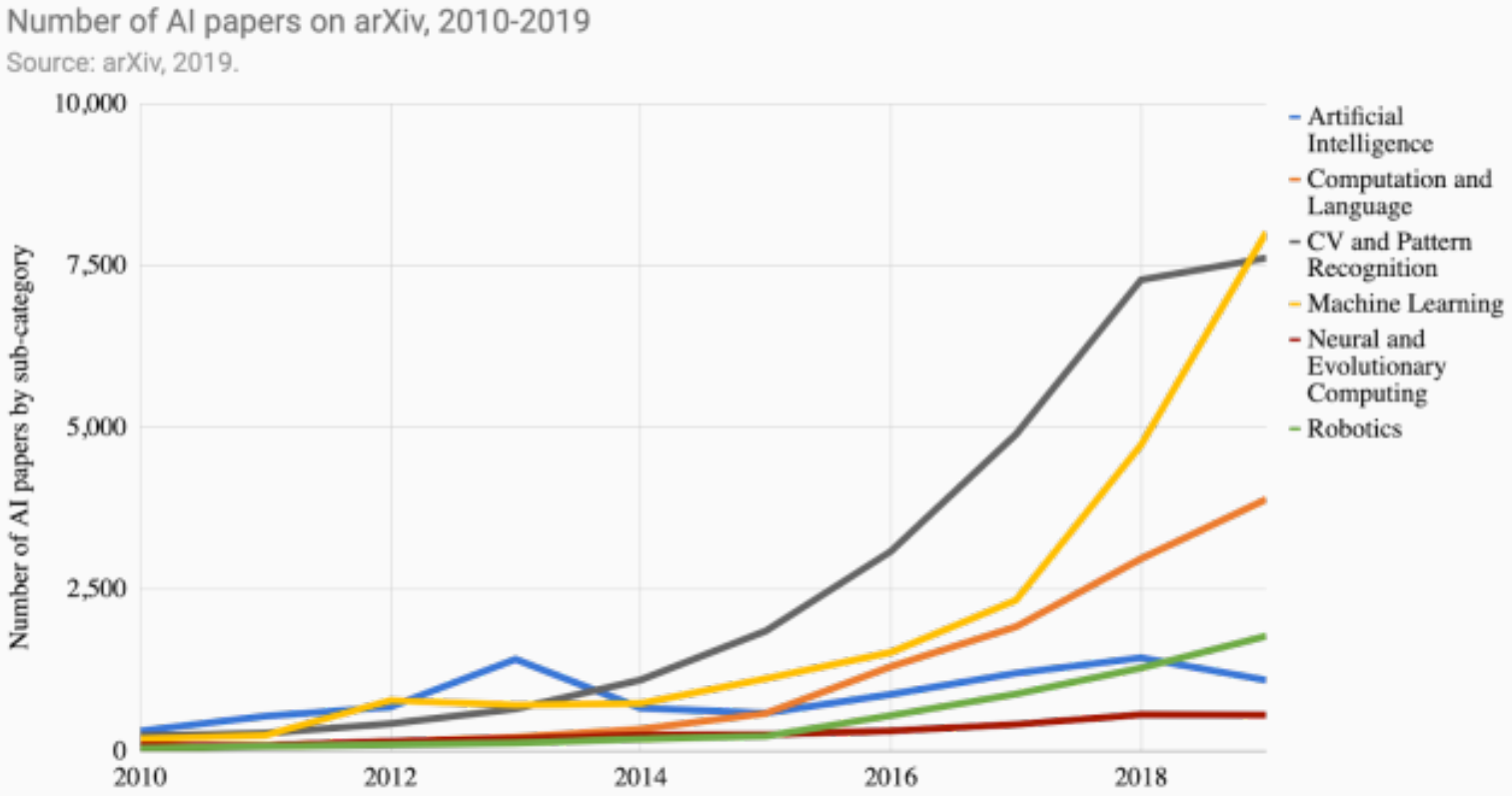
Conclusions

Experimentation is hard:

- Money for value
- Interpretability is difficult
- Catastrophic forgetting is hard to overcome
- We are far from real world QA systems
- Ideal laboratory settings are far from real world use cases

Related Challenges

Literature grows quickly



Artificial Intelligence Index Report 2019, Stanford University

Trend towards Deep Learning

- Large code base
- Programming GPUs
- Experiments are costly
- Cloud computing
- Results interpretation not always clear

Thank you!

Questions?

Guillermo Echegoyen

J. Deriu, A. Rodrigo, A. Otegi, G. Echegoyen, S. Rosset, E. Agirre, and M. Cieliebak.
Survey on evaluation methods for dialogue systems.
Artificial Intelligence Review, 2020.

G. Echegoyen, A. Rodrigo, and A. Peñas.
Cross-lingual training for multiple-choice question answering.
Procesamiento del Lenguaje Natural, 64, 2020.

 A. Peñas, M. Veron, C. Pradel, A. Otegi, G. Echegoyen, and A. Rodrigo.
Continuous learning for question answering.
In *Tenth International Workshop on Spoken Dialogue Systems Technology (IWSDS)*, 2019.

 M. Veron, A. Penas, G. Echegoyen, S. Banerjee, S. Ghannay, and S. Rosset.
A cooking knowledge graph and benchmark for question answering evaluation in lifelong learning scenarios.

In *International Conference on Applications of Natural Language to Information Systems*, pages 94–101. Springer, Cham, 2020.