Neuro-Symbolic Representations for IR
1.1 – Purpose of this Tutorial

Laura Dietz

Neuro-Symbolic Representations...

- Combine strengths of neural networks and symbolic methods
- Neural: continuous representations to capture noisy concepts
- Symbol: capture discrete aspects of the task
- Symbols to help
  - ▶ resolve ambiguity in text, such as entity annotations
  - ▶ draw on structured external knowledge, such as knowledge graphs
  - ▶ derive rich entity representatinos and topically relevant connections
  - ▶ reasoning with compositional logic, such as First-order logic of predicates
- Goal: Combining different types of knowledge in a unified framework

.. Representations for Information Retrieval

- Symbolic entity representations can be incorporated into IR models
- Capture semantic and symbolic information (ex: entities)
- Text and symbolic entity representations can enhance:
  - ▶ Query rewriting: semantically related terms
  - ▶ Semantic matching: understand underlying meaning
  - ▶ Document ranking: what is more relevant
  - ▶ Background Info: connect to Knowledge graphs
  - ▶ Complex information: connect pieces of relevant information
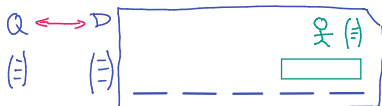
chatGPT criticism: Faithfulness

ChatGPT hallucinates facts. You don't know which are correct.

LA   What is the usual time difference between London and Boston?

The usual time difference between London and Boston depends on the time of year, as both cities observe daylight saving time, but they begin and end at different dates.

During standard time:

* London is in GMT (Greenwich Mean Time)
* Boston is in EST (Eastern Standard Time)
  There is a 5-hour difference, with Boston being 5 hours behind London.

During daylight saving time:

* London is in BST (British Summer Time)
* Boston is in EDT (Eastern Daylight Time)
  There is a 4-hour difference, with Boston being 4 hours behind London.

Keep in mind that the dates for daylight saving time may change, so it's always good to double-check the current time difference between the two cities.

# chatGPT criticism: Faithfulness

ChatGPT hallucinates facts. You don't know which are correct.

Approaches
- Retrieval-augmented Generation: Retrieve provenance first
- Cross-check with complementary information sources
- Verify with manually-checked Knowledge Graph
- Ask for its Chain-of-thoughts, to check each.

Verification: break down information into smaller facts
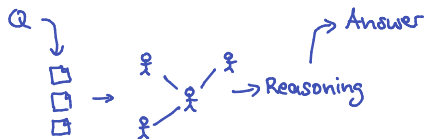$\rightarrow$Symbols
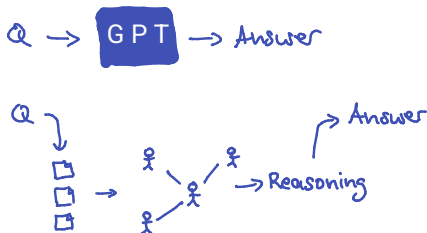
Current approach: More training data (Lambda)

Mix Generative and Discriminative LLM

- Key goals: quality, safety, and groundedness
- Grounded in an external (verified) knowledge resource

1. Ask crowd workers for information seeking queries
2. Generate search query for retrieval
3. Generate answer from retrieved documents
4. Assessors judge how close response follows the source
5. Train with assessor's feedback

Thoppilan 2022. "Lamda: Language models for dialog applications"

Smarter Approach?

Maybe we have smarter ideas?

.oO(Maybe we figure it out today?)
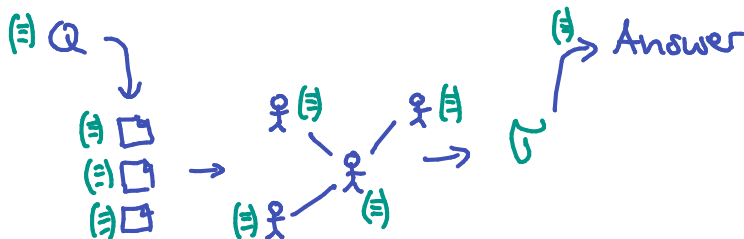
Why not symbolic?

Current state-of-the-art in ad hoc text retrieval
"Only" neural text representations

Why not neuro-symbolic?
- Is it not helping?
- Is it too complicated to implement?
- Did we not yet figure out how to reap benefits?

.o0(Maybe we explore some new ideas today!)

# BERT struggles to understand Entities

Common thinking: BERT can do everything!

– including understanding entities

But: Empirical evidence suggesting that BERT (Transformer-based) embeddings struggle to represent entities. Otherwise we would not need to put as much work in Entity Ranking.

… more on that later.

We know GPT is struggling, but not exactly what the issue is.

Do Entities help?

- In 2014–2019: Entity L2R features improve text ranking
- In 2019–2023: Entity features are not helping neural rankers
- 2022: Significant improvements on LLM-based entity ranking
- Today: Need to reconsider ad hoc text ranking with new "prescription-strength" entity representations
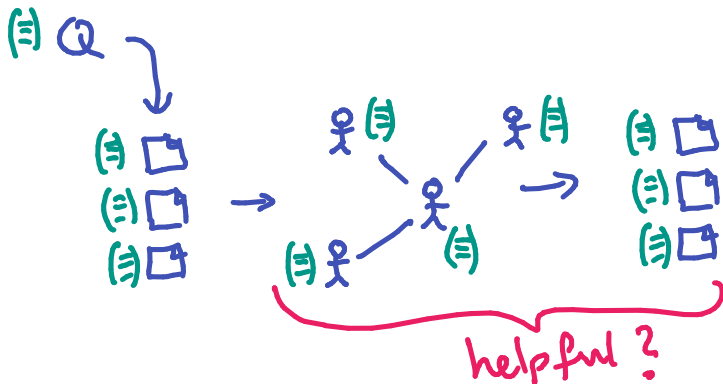
Conversational Search & Assistance

To drive a conversation,

- need to track important concepts (=entities) in the conversation
- reason which concepts are relevantly related
- integrate textual provenance

Prime application for Neuro-Symbolic IR!
...more on that later.

This tutorial

- Overview of progress made on Neuro-Symbolic Representations
- Cover entities, KGs, graph approaches
- Spectrum between purely neural and purely symbolic

Goals:

- Trigger new ideas for new papers
- Spur exciting discussions
- "Stick our heads together" to figure out better approaches

https://github.com/laura-dietz/neurosymbolic-representations-for-IR/

Tutorial Timetable

https://github.com/laura-dietz/neurosymbolic-representations-for-IR/SIGIR23/