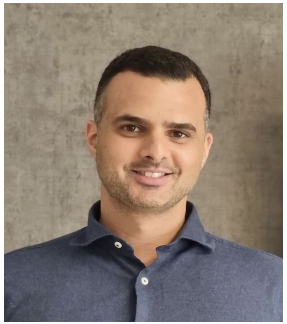
Life is a Circus and We are the Clowns :
Automatically Finding Analogies between Situations and Processes



Oren Sultan

@oren_sultan

oren.sultan@mail.huji.ac.il

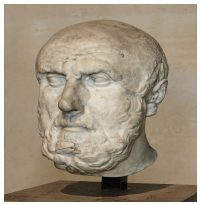
Prof. Dafna Shahaf

@HyadataLab

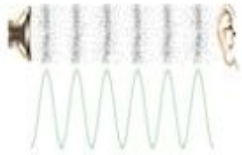
dshahaf@cs.huji.ac.il

Analogies in **human cognition**

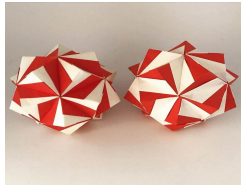
- Analogy-making is a **central** part of **Human Cognition** *(Minsky, 1988; Hofstadter and Sander, 2013; Holyoak, 1984)*
  - Abstract information.
  - Inventions throughout history.

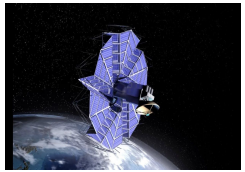


Chrysippus



NASA





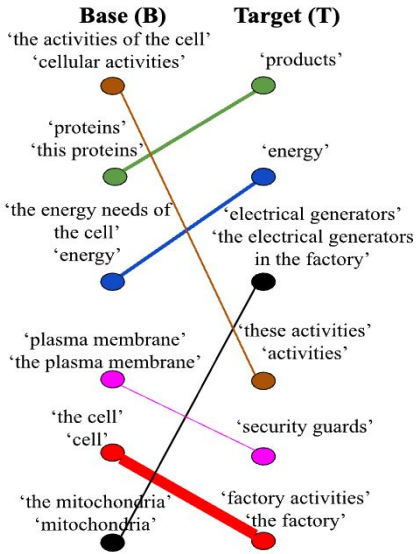
Analogies in **Artificial Intelligence (AI)**

- Analogies are **essentials** for **Artificial Intelligence (AI)** *(Mitchell, 2021)*
  - Key to non-brittle AI systems that can adapt to new domains.
  - Form humanlike concepts and abstractions.
- Analogies in **Natural Language Processing** (**NLP**)
  - Most works focused on word analogies - “a to b is like c to d” *(Mikolov, 2013)*
- **Our focus**: Analogies between **situations** and **processes**
  - Structure Mapping Engine (SME) *(Gentner, 1983; Falkenhainer, 1989; Turney, 2008; Forbus, 2011)*
    - **Input**: two domains
    - **Goal**: map objects from **base** to **target** according to **relational structure** rather than **object attributes**.
    - **Problem**: the input in a highly structured language.

Our work – analogies between natural language procedural texts

Base: Animal Cell
The plasma membrane encloses the animal cell. It controls the movement of materials into and out of the cell. The Nucleus controls the activities of the cell. These cellular activities require energy. The Mitochondria extract energy from food molecules to provide the energy needs of the cell. Animal cells must also synthesize a variety of proteins and other organic molecules necessary for growth and repair. Ribosomes produce these proteins. The cell may use these proteins or move them out of the cell for use in other cells. To move organic molecules, the cell contains a complex system of membranes that create channels within the cell. This system of membranes is called the endoplasmic reticulum.

Target: Factory
Security guards monitor the doors of the factory. They control the movement of people into and out of the factory. Factory activities may be coordinated by a control center. These activities require energy. The electrical generators in the factory provide energy. The factory synthesizes products from raw materials using machines. The factory has hallways to move products through it.



Our Method – Analogous Matching Algorithm

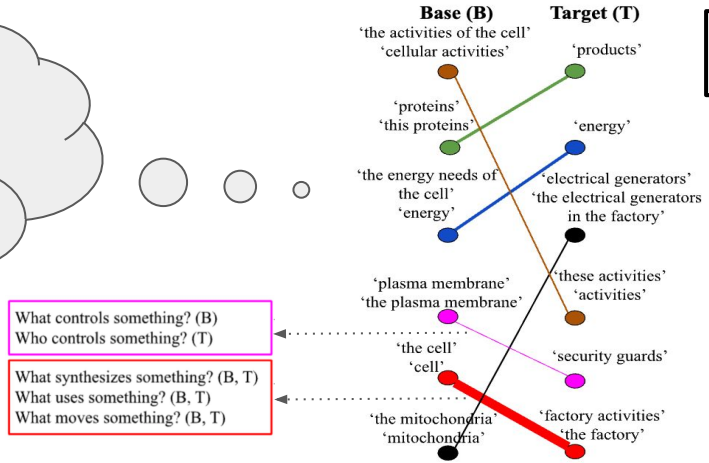
Text processing → Structure Extraction → Clustering Entities → Find Mappings

The Main Idea

- **Q**: How can we know that **entities** in the domains **play similar roles**?

- **QA-SRL** model *(FitzGerald, 2018)*
  - **Input:** A sentence. **Output**: questions and answers about the sentence.
  - The **answers** form the **entities**.
  - **Similar questions** between the domains, indicate that the entities may **play similar roles**.

Find a mapping that **maximizes** the **similarity** between the **questions** of the mapped **objects.**



Find Mappings by Questions (FMQ)

Find Mappings

- We find a mapping that **maximizes** the **similarity** between the **questions** of the mapped **objects**.
- More details in the paper.
- We call our method: **Find Mappings by Questions (FMQ)**.

Results

- **Experiment 1: Mining Analogies**
  - We show that we can use our method (**FMQ**) to mine analogies from the **ProPara** dataset.
  - SBERT was able to find almost only paragraphs on the **same topic**,
    our method has found many **close** and **far** analogies.

- **Experiment 2: Evaluating the Mappings**
  - We show that our method produce the correct mapping solution for:
    - Analogous pairs of paragraphs from **ProPara**.
    - Analogous pairs of **Stories** from cognitive-psychology.

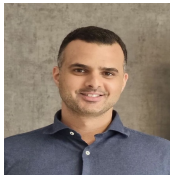
- **Experiment 3: Robustness to Paraphrases**
  - We show that our method is **robust to paraphrasing** the input texts.
    - **Automatic Paraphrases**: using **wordtune**. AI21labs
    - **Responses to the same prompt**: different authors writing on the same topic.

Conclusions

- Analogies are important for **humans** and **AI**.
- We explored analogies between **procedural texts**.
- We developed a method that find a **mapping** based on **relational similarity**.

- Our method was able to **mine different type of analogies** (in contrast to **SBERT**).
- Our method produced the correct mappings on both **ProPara** and the **Stories**.
- We showed our method is **robust to paraphrasing**.

- **Code repository**: https://github.com/orensul/analogies_mining
- Looking forward to seeing you in Abu Dhabi! 🇦🇪😎



Thank You!

@oren_sultan

oren.sultan@mail.huji.ac.il

@HyadataLab

dshahaf@cs.huji.ac.il