Extraction of Metaphoric Analogies from Literary Texts:

Task Formulation, Dataset Construction, and Evaluation

Joanne Boisson[1], Zara Siddique[1], Hsuvas Borkakoty[1], Dimosthenis Antypas[1], Luis Espinosa-Anke[1,2] and Jose Camacho-Collados[1]

[1]Cardiff NLP, School of Computer Science and Informatics, Cardiff University, [2]Amplyfi

4-term metaphoric analogies

Metaphoric mapping

- **Target domain:** semantic domain of the concepts **used literally** in the analogy.
- **Source domain** : semantic domain of the concepts **used metaphorically**.

Structured metaphoric analogy
T1 is to T2 what S1 is to S2

- **Explicit terms** occur in the text.
- **Implicit terms** do not occur in the text, but are conveyed by the reader for understanding the analogy.

i : number of implicit terms
<> : implicit terms

Metaphoric analogies in literary texts	T1	T2	S1	S2	i
Books are like **imprisoned souls** till someone takes them down from a **shelf** and frees them. (S. Butler)	books	shelf	imprisoned souls	<prison>	1
Money is the **mother's milk** of **politics**. (J. Unruh)	money	politics	mother's milk	<baby>	1
An election is coming. Universal peace is declared, and the **foxes** have a sincere interest in prolonging the lives of the **poultry**. (G. Eliot)	<candidates>	<voters>	foxes	poultry	2

- **Task validation:** 5 annotators with background in linguistics or metaphor studies extract analogical frames from 20 short texts given the term T1.
- Average pairwise inter-annotator agreement scores with a **Cohen's Kappa score of 0.68 for the frames containing explicit terms**
  - **metric** : lemmatized head-noun match

Dataset construction: 204 short texts are selected from existing collections of literary metaphors and labelled

Task

Input :
- a short text
- a term T1, T2, S1 or S2

Instruction :
- Extract the other explicit terms forming the 4-term metaphor
- Generate eventual missing implicit terms

Output :
The structured metaphoric analogy : values of the 4 frames **T1**, **T2**, **S1** and **S2**

Experiments

7-shot in-context learning with 5 models
- Llama-3, Mixtral 8*7 & Mixtral 8*22
- GPT-3.5 & GPT-4

Dataset instances are tested with
- Each of the 4 frames given as input
- 3 different example sets in the prompt

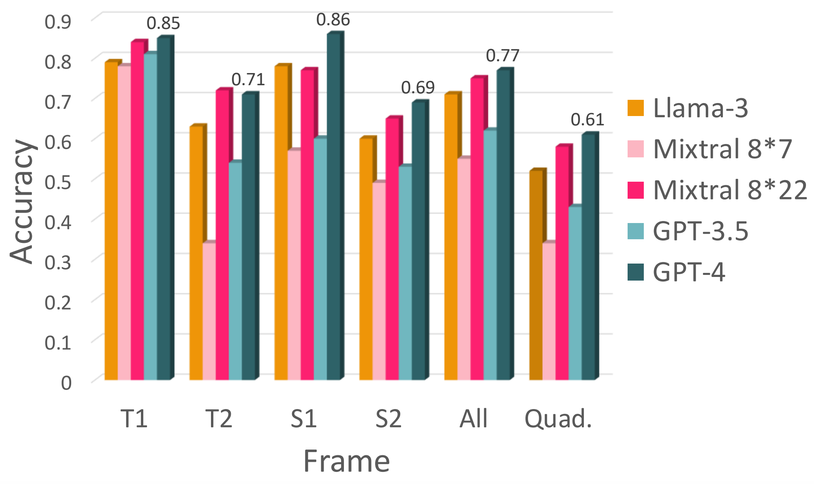
Evaluation of the extracted explicit terms

Metric: Lemmatized head noun

All : correct when any frame T1 T2, S1 or S2 is correctly extracted (669 distinct explicit terms)

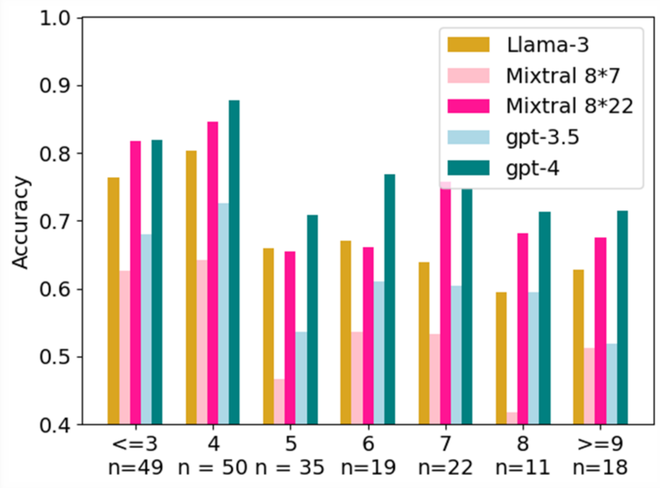
Quad. : correct when all the explicit terms of a quadruple are correctly extracted (204 distinct quadruples)

Accuracy results for explicit terms split by output frame



n : number of short text in each bucket

Accuracy results for explicit terms split by number of nouns in the input text



Evaluation of the generated implicit terms

Manual evaluation: 2 annotators rate the generated implicit terms of 50 different examples, where explicit terms were extracted correctly.

0 : incorrect
1 : imperfect
2 : very good

Example : **Written laws** are like **spiders' webs**, and will, like them, only entangle and hold **the poor and weak**.

Explicit terms: T1 : written laws; S1 : spiders'webs; T2 : the poor and weak

Suggested implicit term S2 : <weak insect>

Model	Generated S2 terms	Rating
Mixtral 8*22	spiders	0
	entangle and hold	0
GPT-4	entangled beings	1
	small insects	2

- **Sentence score:** average rating of all generated terms
- **Model score:** average sentence score

Score of the best two models:

Model	Total	Deduplicated	Score
Mixtral 8*22	356	153	0.75
GPT-4	369	188	1.21

- **Total** : total number of evaluated terms
- **Deduplicated :** number of distinct generated words

Conclusion

- Language models can become valuable tools for the conversion of unstructured metaphors to structured analogical concept mapping.
- The performance of the best models analyzed, namely GPT-4 and Mixtral 22*8 are in line with human annotators.

Possible extensions of the analogy framework
- n to n concepts mapping extraction
- Relation extraction

Questions for future work
- How robust are the models for analogy extraction on open longer texts?
- What are the type of metaphoric analogies that the models struggle the most to extract?

Repository