

MultiZebraLogic: A Multilingual Logical Reasoning Benchmark

Sofie H. Bruun, Dan Saattrup Smart

The Alexandra Institute

Zebra puzzles require multi-step reasoning

Automatic puzzle generation w. input from native speakers

Phrases e.g. "the baker", "is a baker", "is not a baker" + grammar

9 Germanic languages, built for scalability

Dataset of 2x3 and 4x5 puzzles

Red herrings increase difficulty

To be implemented in EuroEval

	Attributes		
Objects	object_1	teacher	cat
	object_2	nurse	dog

A row of houses have numbers 1 to 2 from left to right.

In each house lives a person with unique attributes in each of the following categories:

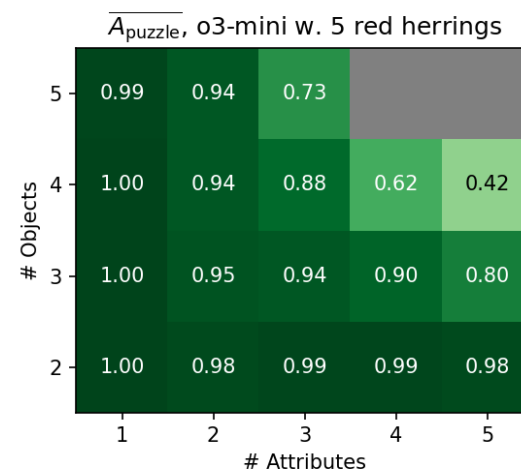
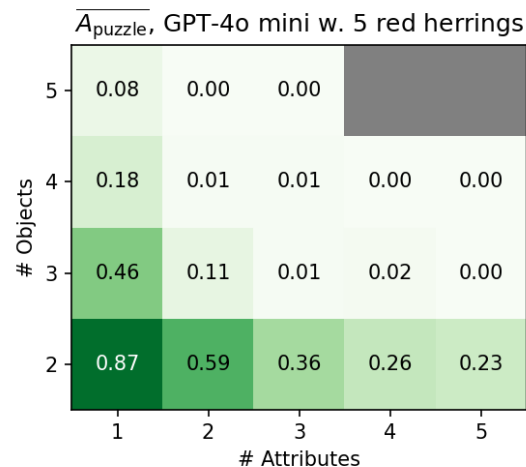
Jobs: nurse and teacher.

Pets: cat and dog.

We also know the following:

1. The cat owner lives to the left of the nurse.

Who has which attributes and lives in which house?



TrustLLM



Funded by the European Union