Coordinates from Context: Using LLMs to Ground Complex Location References

Currently on the job market!

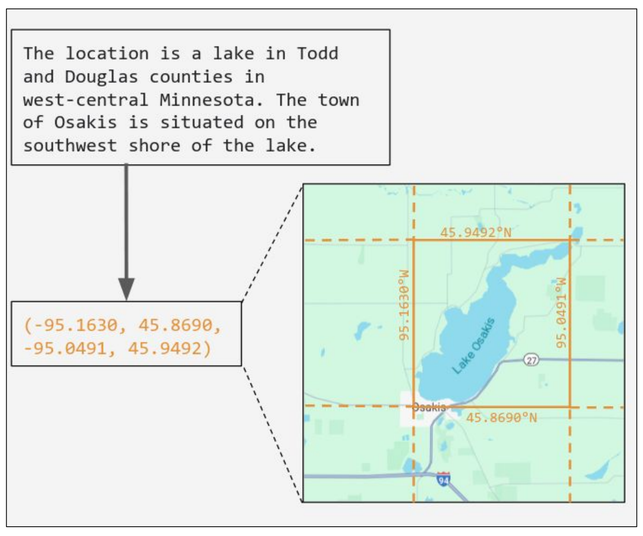
Tessa Masis (tmasis.github.io/) and Brendan O'Connor

University of Massachusetts Amherst

Compositional Locations

"Along the border of Todd and Douglas counties, the lake is…"

How can we ground locations that are indirectly referred to as compositional descriptions of other places?
They may not have an entry in a geographic database, which most geocoding methods rely on



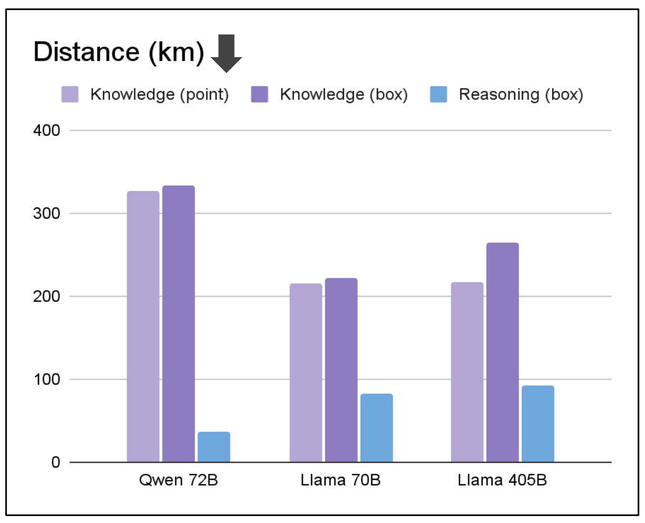
Idea: given a compositional location description, ground it to an actual geographic location by predicting its bounding box

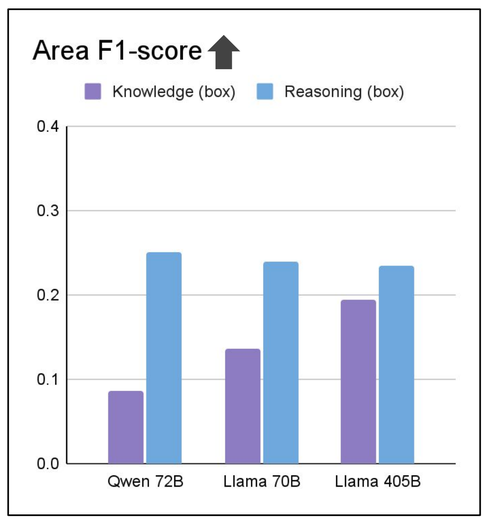
Do LLMs Know Geographic Bounding Boxes?

Knowledge: predict the center coordinates/bounding box of a location *identified unambiguously by name*

Hotel Astoria, in Denmark → (55.6734, 12.5631)

Hotel Astoria, in Denmark → (12.5628, 55.6730, 12.5634, 55.6738)





Reasoning: predict the bounding box of an *unidentified location described in relation to other toponyms*

The location is a design hotel next to the Central Station in Copenhagen, Denmark. Central Station has a longitude of… Copenhagen has a longitude of… → (12.5628, 55.6730, 12.5634, 55.6738)

We find that off-the-shelf LLMs have stronger geospatial reasoning skills than geospatial knowledge!

Data

The location is a lake in Todd and Douglas counties in west-central Minnesota. → (-95.163, 45.869, -95.049, 45.949)

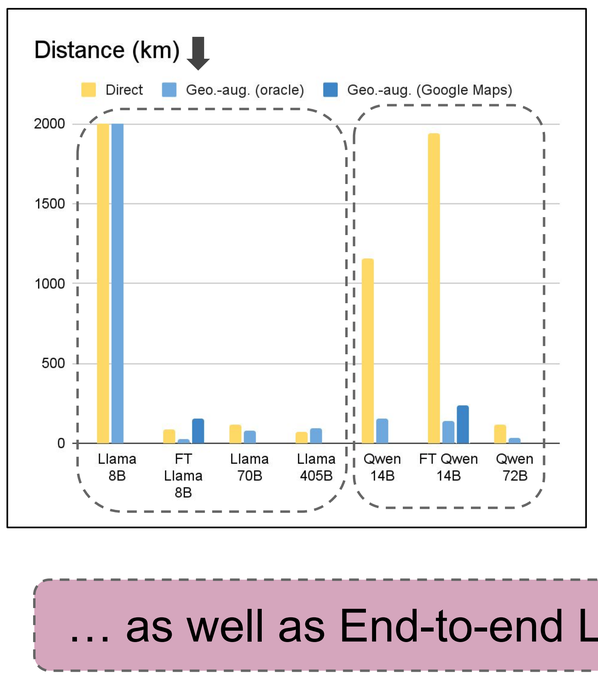
GeoCoDe dataset (Laparra & Bethard, 2020)
Location descriptions from Wikipedia articles, linked to their bounding boxes from the OpenStreetMap geodatabase

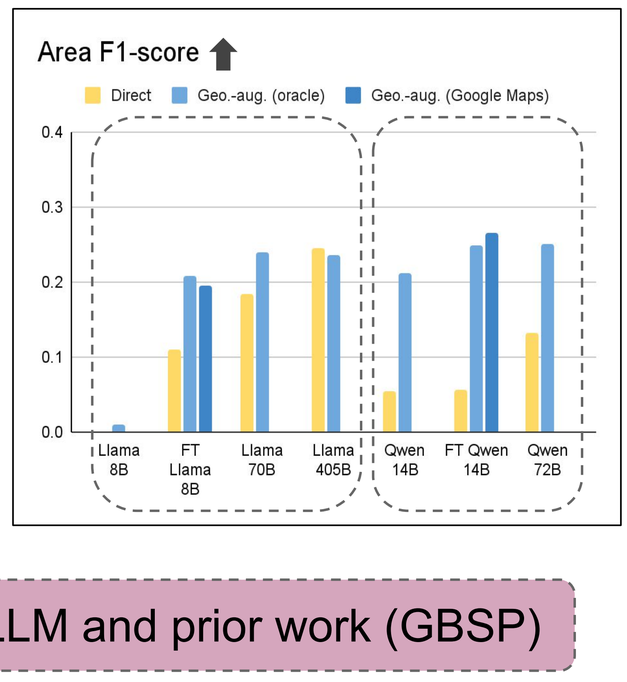
Tool-Augmented Geocoding

The `recaller`, a traditional geoparsing tool (e.g. Google Maps Geoparser), retrieves mentioned locations' center coordinates

The `reasoner`, an LLM, uses the location description + mentioned locations' coordinates to generate the described location's bounding box

Geoparser-augmented outperforms no `recaller` (Direct)...





… as well as End-to-end LLM and prior work (GBSP)

Approach	`reasoner`	Coverage (%) ↑	Distance (km) ↓	F1-score ↑
Geo.-aug.	FT Qwen 14B	90.8	**240.9**	**.266**
End-to-end LLM	FT Qwen 14B	**95.9**	627.5	.167
GBSP	Semantic parser	52.8	–	.240

Common Errors from LLMs?

- **Flipped signs**, *e.g. "95.163" instead of "-95.163"*
- **Overreliance on in-context information without reasoning about its relevance** *(e.g. predicting max and min mentioned locations' coordinates as the bounding box)*, resulting in larger & less precise bounding boxes