NUI Galway
OÉ Gaillimh

Digital Enterprise Research Institute

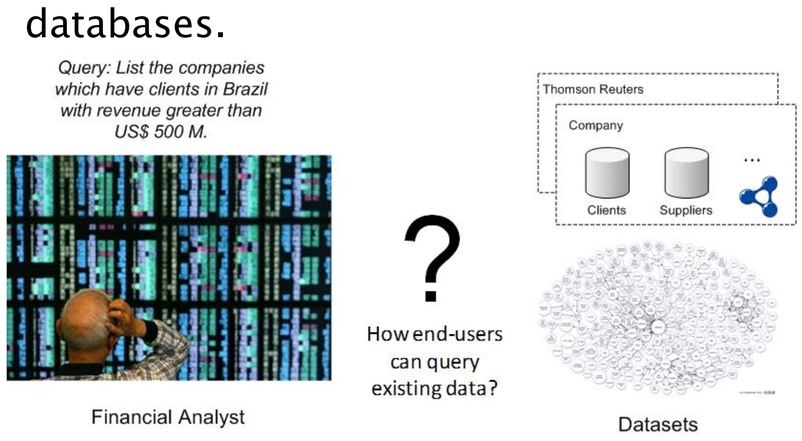
DERI

A Distributional Structured Semantic Space for Querying EAV Databases

André Freitas, Seán O'Riain, Edward Curry

Motivation

- Currently users need to be aware of the structure of datasets in order to query them.
- Natural language queries as an intuitive way to query Entity-Attribute-Value databases.

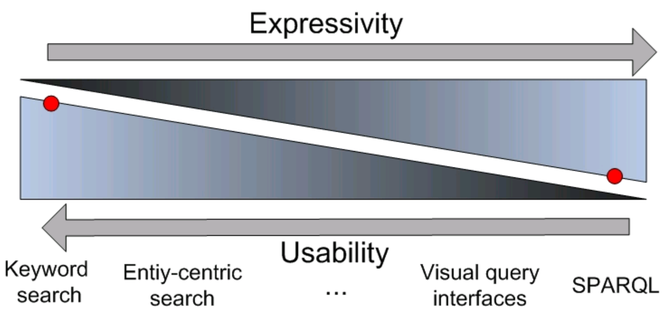


Objective

- A vocabulary independent natural language query mechanism.

Challenges

- How to allow users query data without prior knowledge of its representation?
- How to address the expressivity-usability trade-off of existing approaches?



Proposed Approach

- Combines of 3 elements:
  - **Entity search.**
  - **Distributional semantics: semantic relatedness**
  - **Principled compositional semantics: spreading activation search.**
- Formalized as a *distributional structured vector space model* (T-Space).

Results

Evaluated using the QALD Workshop query set containing 100 natural language queries over DBpedia.

- **The approach was capable of answering 83% of the queries.**
- **Avg. precision = 0.61, avg. recall = 0.86, mean reciprocal rank = 0.49.**
- The distributional-compositional model:
  - Showed high discrimination for node selection.
  - Deliveries a precise and flexible semantic matching between query and vocabulary terms.

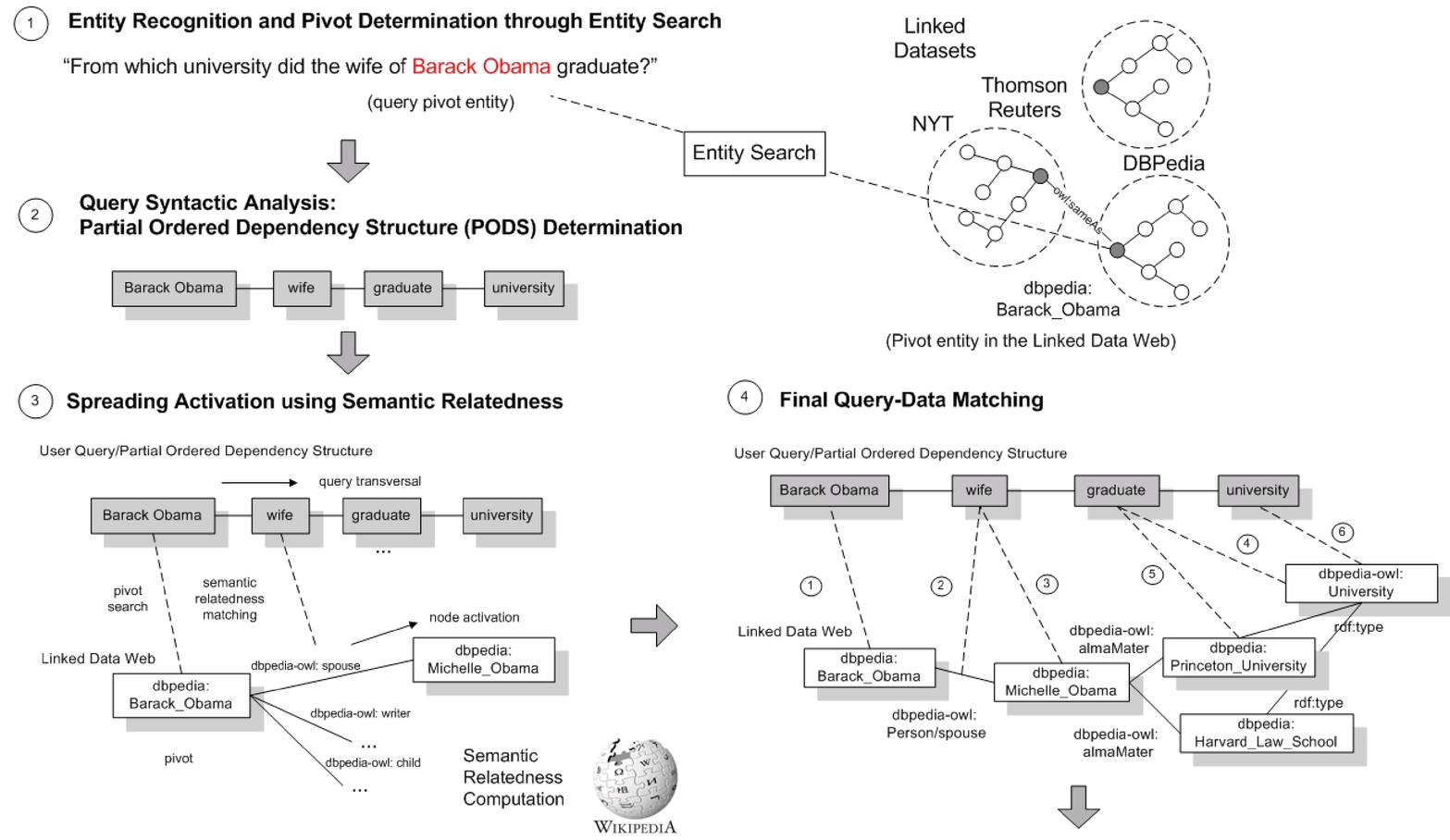
Acknowledgements

This work has been funded by Science Foundation Ireland under Grant No. SFI/08/CE/I1380 (Lion-2).

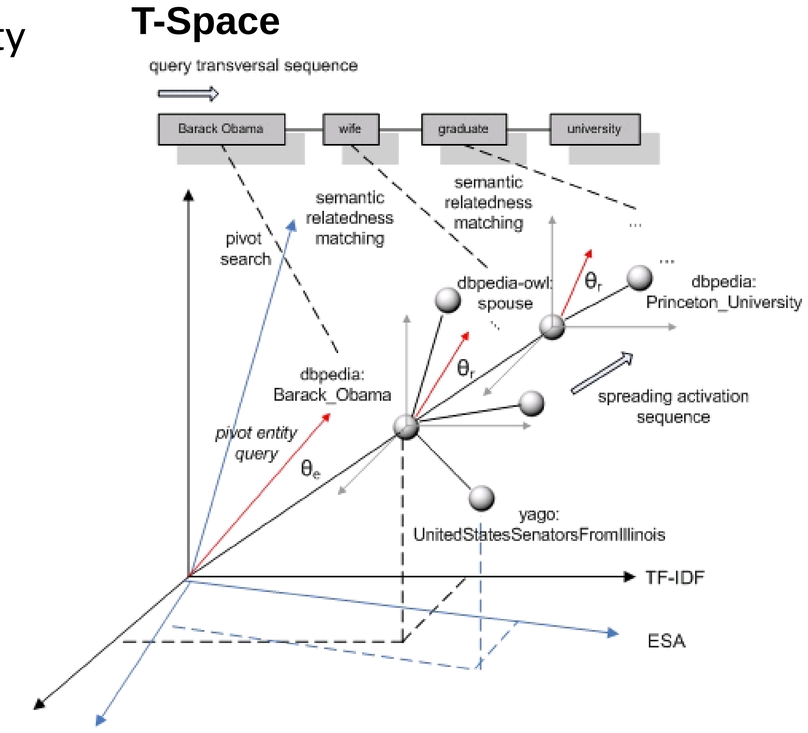
Querying the Linked Data Web

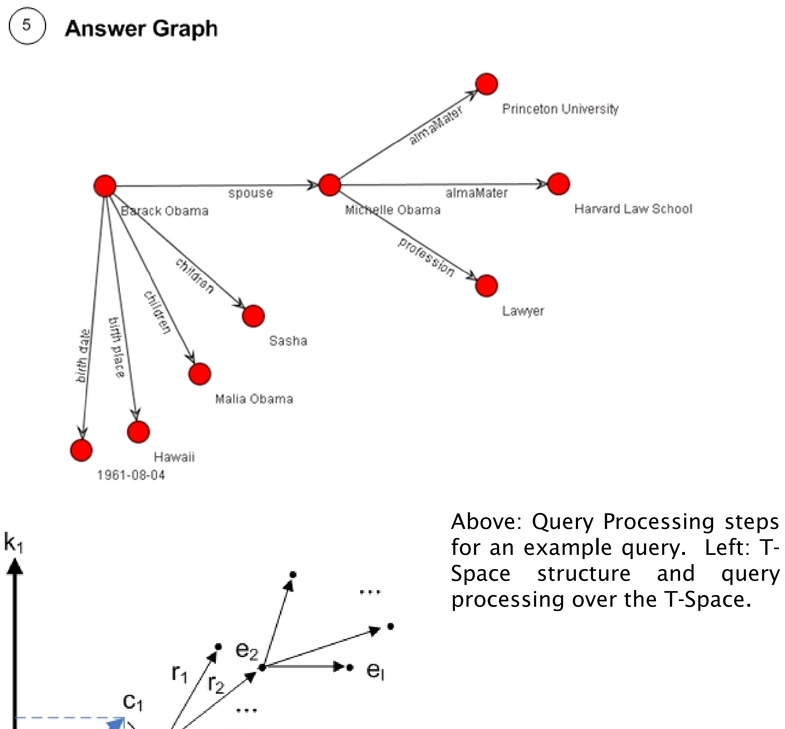
Example query: "From which university did the wife of Barack Obama graduate?"

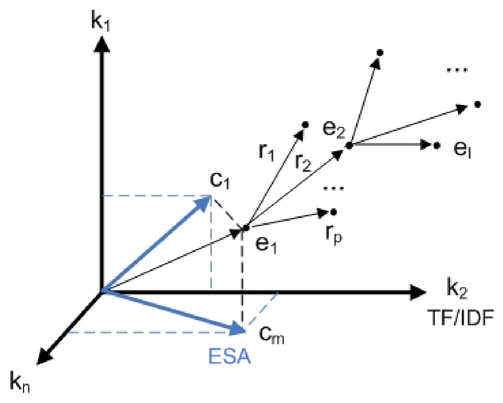
Treo



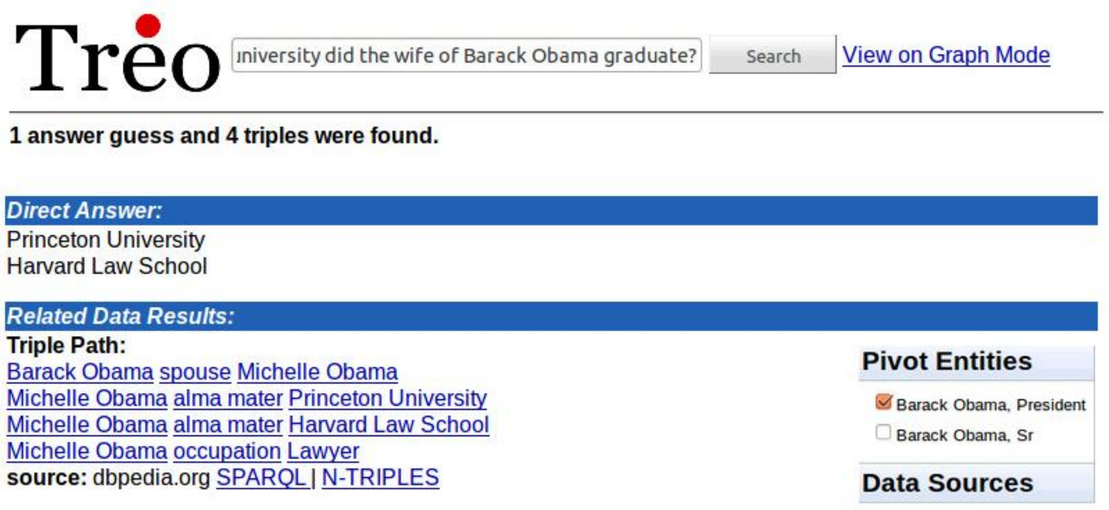
T-Space







Above: Query Processing steps for an example query. Left: T-Space structure and query processing over the T-Space.



References

A. Freitas, E. Curry, J. G. Oliveira, S. O'Riain, A Distributional Structured Semantic Space for Querying RDF Graph Data. International Journal of Semantic Computing (IJSC), 2012.

A. Freitas, J. G. Oliveira, S. O'Riain, E. Curry, J. C. Pereira da Silva, Querying Linked Data using Semantic Relatedness: A Vocabulary Independent Approach. In Proceedings of the 16th International Conference on Applications of Natural Language to Information Systems (NLDB), 2011.

SFI Science Foundation Ireland
ENTERPRISE IRELAND
Ireland's EU Structural Funds Programmes 2007 - 2013
Co-funded by the Irish Government and the European Union
EUROPEAN REGIONAL DEVELOPMENT FUND
IRCSET Investing in People and Ideas
Irish Research Council for Science, Engineering and Technology
SEVENTH FRAMEWORK PROGRAMME

Enabling **Networked** Knowledge