Zero-Shot Open Entity Typing as Type-Compatible Grounding

Ben Zhou[1], Daniel Khashabi[2], Chen-Tse Tsai[3], Dan Roth[2]

[1]University of Illinois, Urbana-Champaign, [2]University of Pennsylvania, [3]Bloomberg LP

1.Goal

Given ***mention*** in its ***context***, assign ***a set of types*** that the mention belongs to.

In this work, we extend this goal to type with an **open** type set, without reliance on task specific supervision (**zero-shot**).

2.Motivation

Most previous works focus on using supervision for typing, which often rely on costly annotation, do not generalize to unseen types or out-of-domain data.

Unlike previous systems, we propose an approach that is both **`open`** and **`zero-shot`**.

- **`Open`:** Not limited to a fixed label set
- **`Zero-shot`**: No need for typing-specific annotation

To achieve zero-shot, we equip our model with **`understanding`** of the labels. We propose the notion of **`type-compatible`**, which transforms the typing problem to approximating type-compatibility, by mapping an input mention to a set of type-compatible entities, then inferring their types.

- **`Understanding`**: Each type is implicitly defined by entities that share that type.
- **`Type-compatible`:** Two mentions are type-compatible if they share at least one type.

3.Model

In order to compute type-compatibility between mention and entities, we propose a notion called **context-consistency.**

- **Context-consistency:** two mentions are context-consistent if they are interchangeable in some context. In other words, it represents the contextual similarity between two mentions.

We approximate the **type-compatibility** between a mention and an entity (represented by its mentions) through computing **context-consistency**.

A significant *advantage* of using context-consistency is that it relies mostly on context, which overcomes the issue of *polysemy* or *rare* mention surface forms. *(see Figure 1)*

Acknowledgements

This work was supported by the US Defense Advanced Research Projects Agency (DARPA), Google and the Allen Institute for Artificial Intelligence.

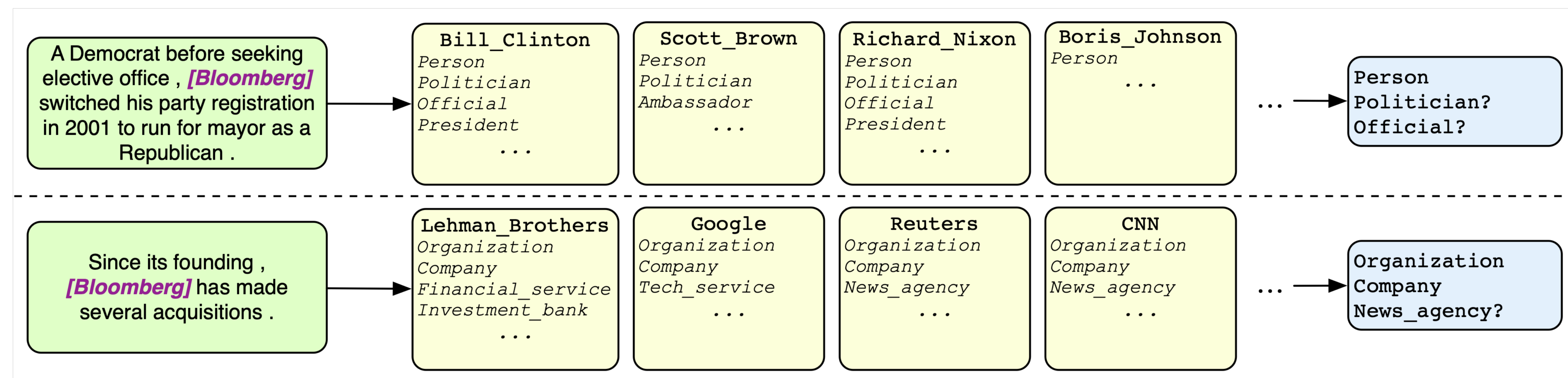


Figure 1: Example inputs, corresponding compatible entities and output

4.Details

We use **`ESA`** to generate a relatively large set of initial concepts.

Then we use **`ELMo`** representations to approximate context-consistency between a mention and a concept. A concept's representation is the averaged representations of its mentions in **`WikiLinks`**.

We take a small set of concepts with top context-consistency, then use **`Priors`** to add highly confident concepts based on surface form (if any) to this set.

Then we map the concepts in the small set to target taxonomy using **`FreeBase`** as a proxy (see figure 2), by defining each target type in terms of logical expressions of **`FreeBase`** types.

The final process is a voting based **`Inference`** over types in target taxonomy.

- Ⓐ **`WikiLinks`**: A dataset containing contextual examples for Wikipedia concepts.
- Ⓑ **`ESA`**: Generated by treating each WikiLinks sentence as a document. It enables quick retrieval of topically relevant concepts given a sentence.
- Ⓒ **`ELMo:`** A SOTA contextualized representation, which generates a representation for word in its context.
- Ⓓ **`Priors`**: *P(concept|surface)* table, generated from Wikipedia text and anchor links.
- Ⓔ **`FreeBase:`** A large taxonomy that each Wikipedia concept has a mapping to.
- Ⓕ **`Inference`**: An algorithm that infers the types of the input mention, by taking into account the agreement among the types of the contextually similar concepts.

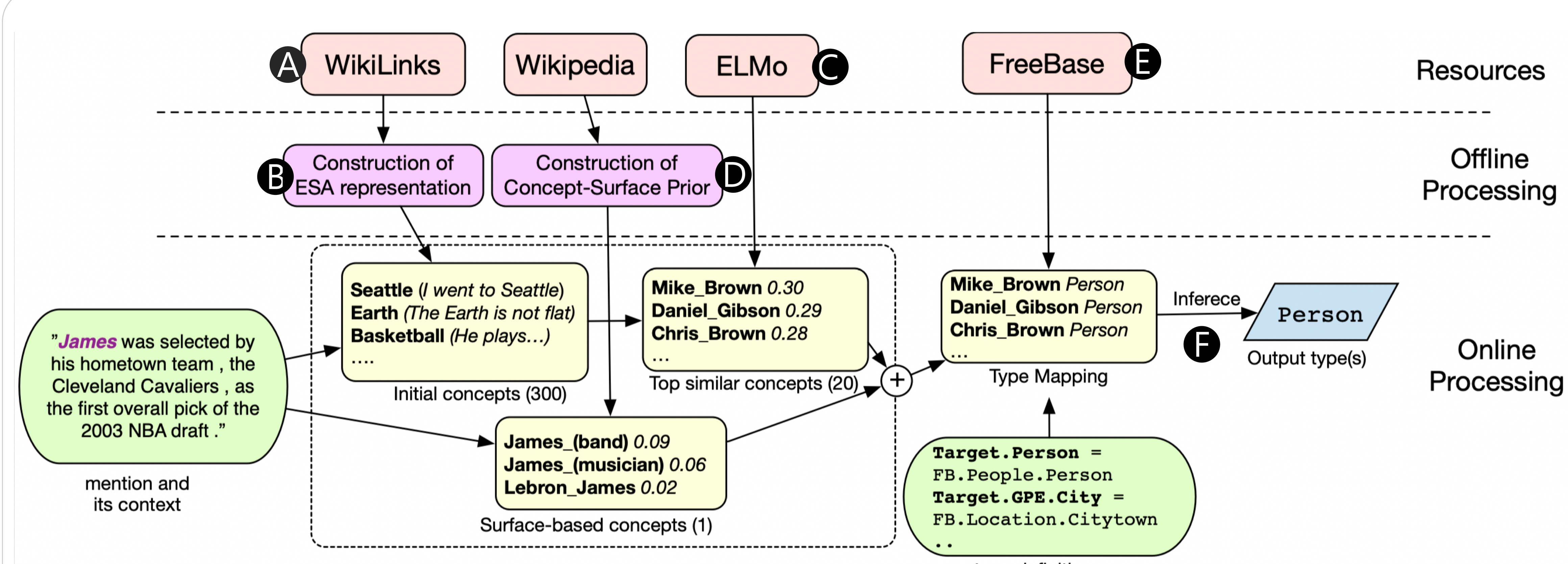
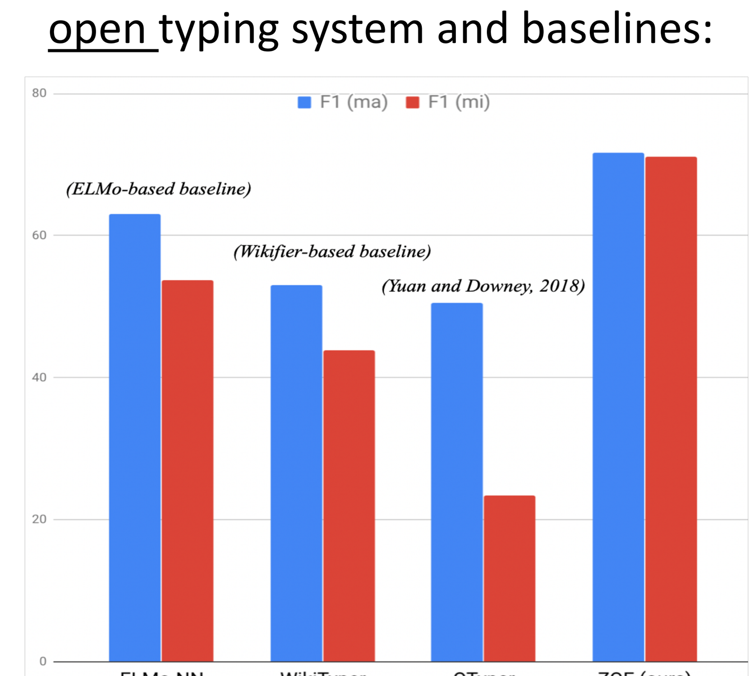


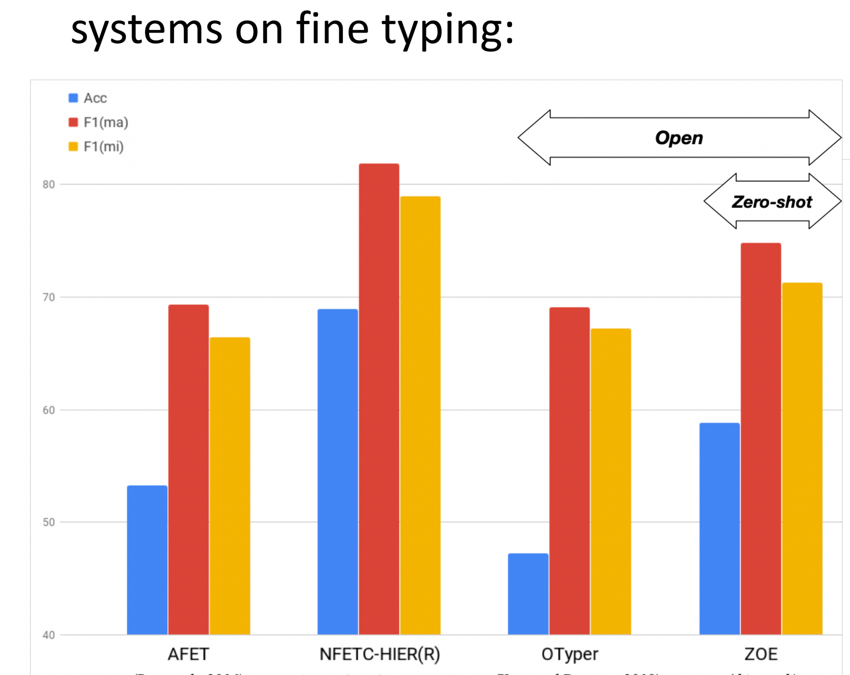
Figure 2: Overview of the system through an example

5.Experiments

Significantly outperforms SOTA open typing system and baselines:

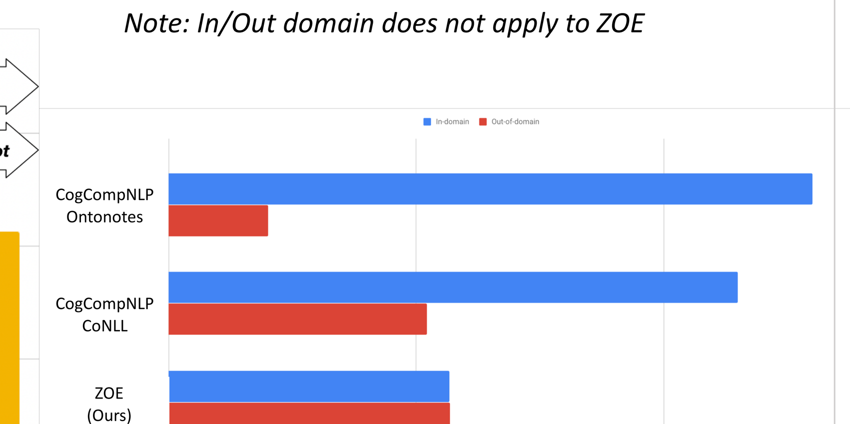


Comparable results with supervised systems on fine typing:



Outperforms supervised system in cross-domain settings

Note: In/Out domain does not apply to ZOE



(more experiments in the paper)