University of BRISTOL

Chapter 17, Speech and Language Processing, 3rd edition draft, Jurafsky & Martin (2020).

6.2. Relation Extraction

Edwin Simpson

Department of Computer Science,

University of Bristol, UK.

bristol.ac.uk

Relations Between Named Entities

- Assume we have already run NER to label the entities:

"Citing high fuel prices, [ORG United Airlines] said [TIME Friday] it has increased fares by [MONEY $6] per round trip on flights to some cities also served by lower-cost carriers. [ORG American Airlines], a unit of [ORG AMR Corp.], immediately matched the move, spokesman [PER Tim Wagner] said. [ORG United], a unit of [ORG UAL Corp.], said the increase took effect [TIME Thursday] and applies to most routes where it competes against discount carriers, such as [LOC Chicago] to [LOC Dallas] and [LOC Denver] to [LOC San Francisco]."

Relations Between Named Entities

- What does the text tell us about how the entities relate to each other?
- **Choose two pairs of related entities and write down their relationship.**

"Citing high fuel prices, [ORG United Airlines] said [TIME Friday] it has increased fares by [MONEY $6] per round trip on flights to some cities also served by lower-cost carriers. [ORG American Airlines], a unit of [ORG AMR Corp.], immediately matched the move, spokesman [PER Tim Wagner] said. [ORG United], a unit of [ORG UAL Corp.], said the increase took effect [TIME Thursday] and applies to most routes where it competes against discount carriers, such as [LOC Chicago] to [LOC Dallas] and [LOC Denver] to [LOC San Francisco]."

Relations Between Named Entities

"Citing high fuel prices, [ORG United Airlines] said [TIME Friday] it has increased fares by [MONEY $6] per round trip on flights to some cities also ser… …rican Airlines], a unit of [ORG … …move, spokesman [… …a unit of [ORG UAL Corp.], s… …sday] and applies to most routes where it competes against discount carriers, such as [LOC Chicago] to [LOC Dallas] and [LOC Denver] to [LOC San Francisco]."

[United Airlines] announced a fare increase on [Friday]

bristol.ac.uk

Relations Between Named Entities

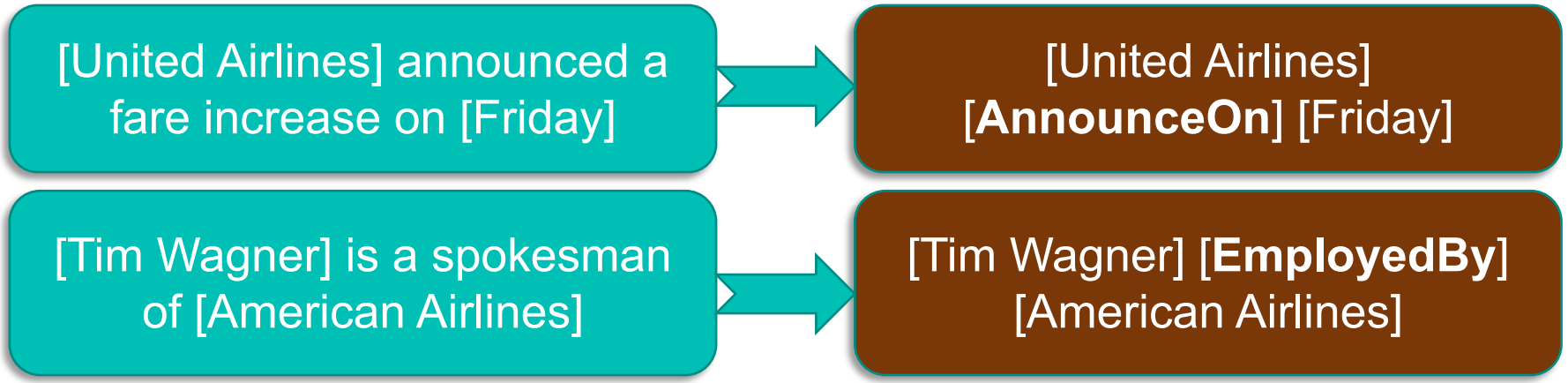
“Citing high fuel prices, [ORG United Airlines] said [TIME Friday] it has increased fares by [MONEY $6] per round trip on flights to some cities also served by lower-cost carriers. [ORG American Airlines], a unit of [ORG AMR Corp.], immediately matched the move, spokesman [PER Tim Wagner] said. [ORG United], a unit of [ORG UAL Corp.], said the increase took effect [TIME Thursday] and applies to most route carriers, such as [LOC Chica o [LOC San Francisco].”

[Tim Wagner] is a spokesman of [American Airlines]

bristol.ac.uk

Relation Extraction

- Let's write the relations to in a formal, structured way.
- We assign each relation to a type:



Relation Extraction (RE)

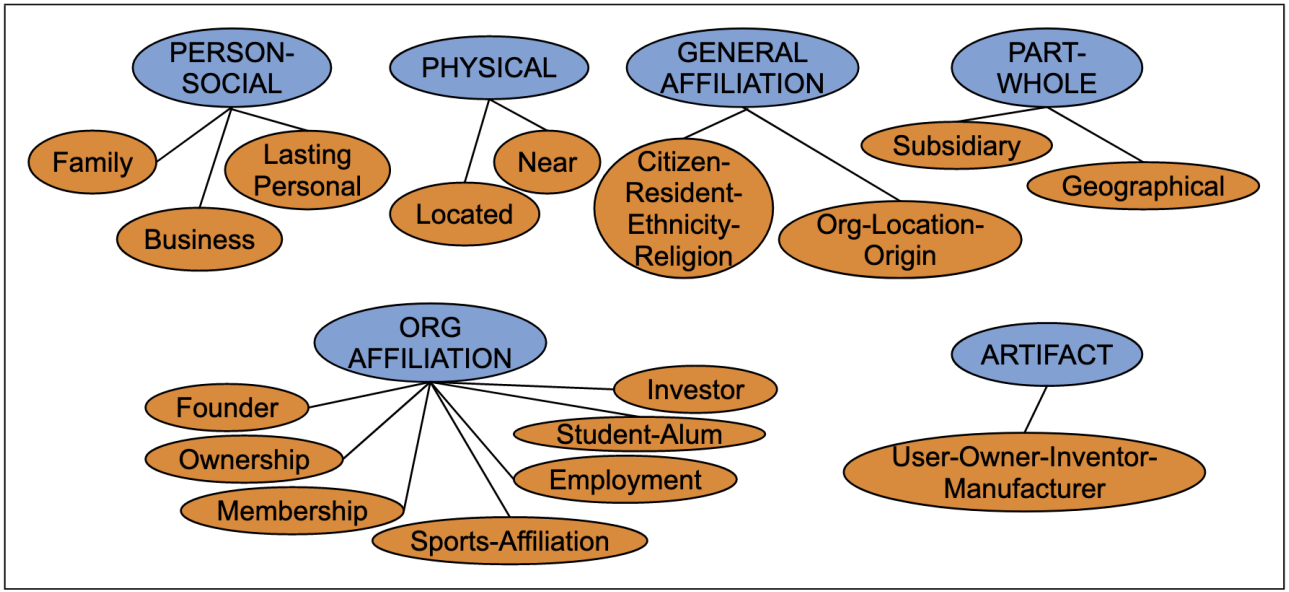
- We just did some manual relation extraction!
- The task of RE is to classify the relation between pairs of named entities in a piece of text.
- For a particular application, we define a set of relation types.

Section 17.1, Speech and Language Processing, 3rd edition draft, Jurafsky & Martin (2020).

ACE Relation Types

Figure 17.1, Speech and Language Processing, 3rd edition draft, Jurafsky & Martin (2020).

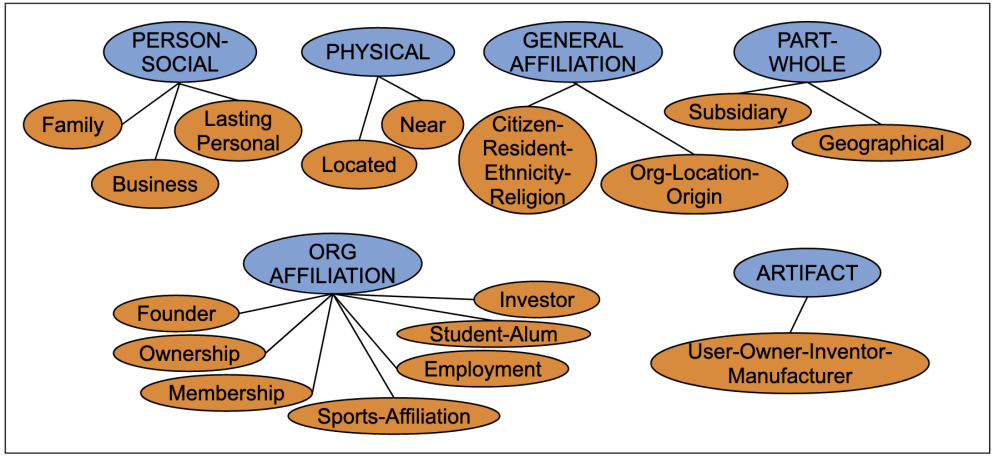
- ACE defines a set of general-purpose relations that are applicable in many scenarios.
- The relations are further grouped into 6 categories.



ACE Relation Types

Figure 17.1, Speech and Language Processing, 3rd edition draft, Jurafsky & Martin (2020).

- Example: “Tim Wagner works for American Airlines” is represented by a relation of type **Org-AFF-Employment**
- This relates named entities of type PER and ORG.
- We can write the relation as a tuple:
  [Tim Wagner]
  [**Org-AFF-Employment**]
  [American Airlines]



Links to Syntax Parsing

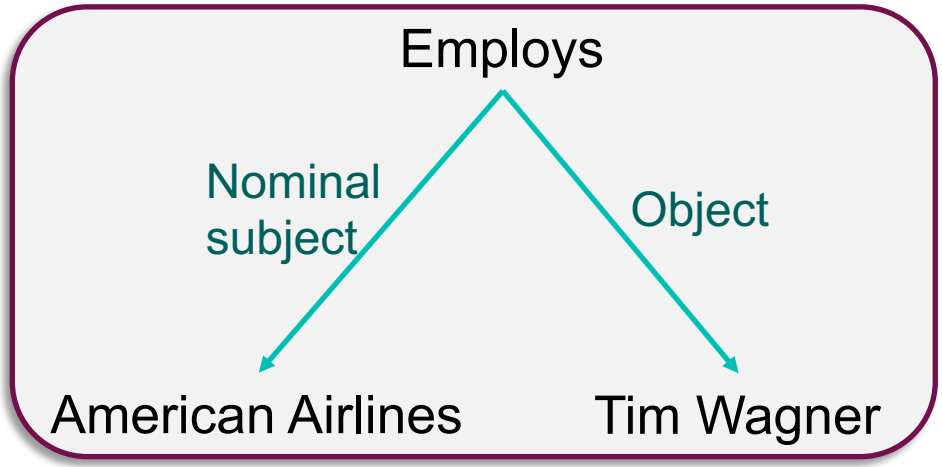
[Tim Wagner] [**Org-AFF-Employment**] [American Airlines]

- This triple is related to something we have seen before…
- We introduced a way to identify verbs, their subject and objects.
- Can you recall which method extracts this information?

Links to Syntax Parsing

[Tim Wagner] [Org-AFF-Employment] [American Airlines]

- This triple is related to something we have seen before…
- We introduced **dependency parsing**, which can identify verbs, their subject and objects.
- More generally, the dependencies can represent predicates and their arguments.
- Dependency parse trees therefore indicate likely relations.



Summary

- The task of Relation Extraction (RE) is to identify how pairs of named entities are related.
- These relations represent fact-like information, such as dates of events or who works for whom.
- Relations are classified into a set of application-specific types.
- Relations can be represented as triples, [named entity 1] [relation type] [named entity 2].
- Dependency parse trees indicate possible relations.