

Hyponymy as a Key to IE.

Rules.

Y such as X $((), X)^* (, \text{and/or}) X$.

Such Y as X

X or other Y

X and other Y

Y including X

Y, especially X $((), X)^* (, \text{and/or}) X$.

X ← hyponym.

Hearst Patterns.

Example occurrences.

X and other Y

temples, monasteries and other important civic buildings.

X or other Y

boisies, wounds, fractures or other injuries

Y such as X

The bow like such as the Rambasa...

Such Y as X

Such authors as Goldsmith, Rindsmoth and Shakespeare.

Y including X

Common law countries including Canada and England.

Y, especially X European countries, especially France, Germany and Spain

Microsoft Research is a part of
Microsoft Corporation.

Neither hyponymy nor hypernymy!

What? Mesonymy & holonymy.

basement of a tall building
part [NN-PL] of [PREP] {the | a} [DET] [JJ/NN] whole [NN]

building's ^{ride} (own) basement
whole [NN-PL]'s past [NN-PL]
~~the~~ ~~the~~ JJ (POS variety).

Disadvantages of Heurst patterns.

- ① Hard to craft rules.
- ② Hard to maintain an updated set of rules.
- ③ Explodes soon.
- ④ Changes with domain ←

{ 66% (Hyponymy)
55% (Mesonymy).

Bootstrapping approach.

↳ something to bootstrap on.

Assumptions.

- ① Not too much annotated data is available for full fledged supervised training
- ② Seed set of annotated data (small one).
↳ do something useful.

↳ semi-supervised.

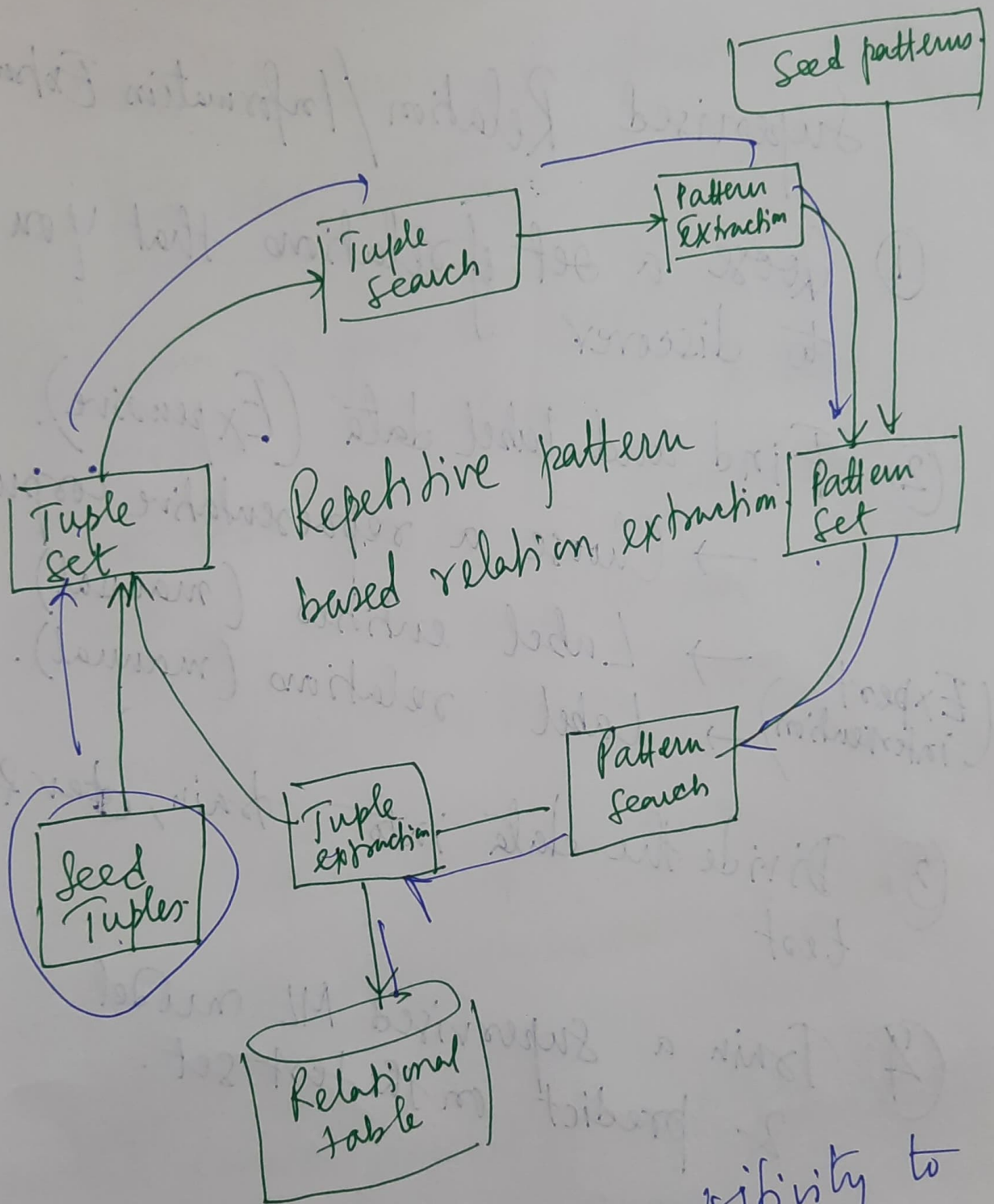
- ✓ Target relation: burial place.
- ✓ seed tuple: [Mask, Train, Elmira].

Google search: "Mask Train" and "Elmira".

"Mask Train is buried in Elmira, NY"
→ X is buried in Y

"The grave of Mask Train is in Elmira"
→ The grave of X is in Y

"Elmira is Mask Train's final resting place."
→ Y is X's final resting place.



Disadvantages? ① Extreme sensitivity to the seed set. ② Too many parameters

→ Probabilistic notion is missing.

→ We are never sure as to how confident we are in the results

Supervised Relation/Information Extraction.

① Choose a set of relations that you wish to discover.

② Find and label data (Expensive).

→ Choose a representative corpus.

→ Label entities (manual)

(Expert intervention) → Label relations (manual).

③ Divide the data into train, dev & test

④ Train a supervised ML model & predict on the test set.

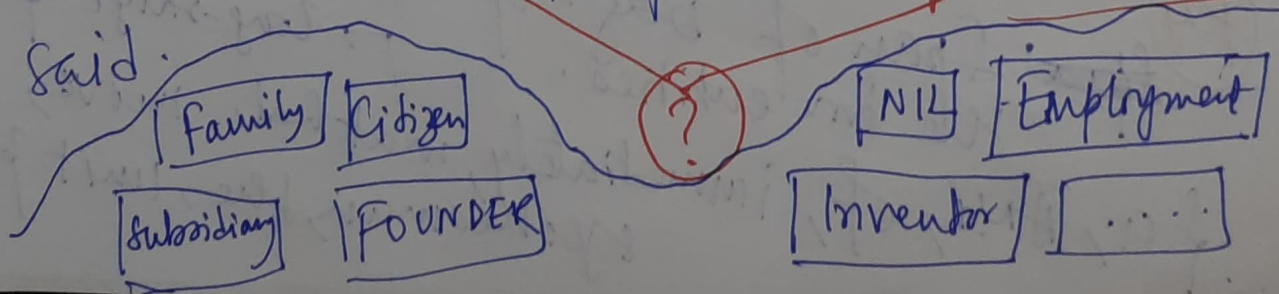
Model SIE as a classification task.

Two step problem.

- ① Given a pair of entities we predict if they are related or not (0/1). — Binary classifier.
- ② If a pair of entities are related, predict the relation type.

→ ~~multi~~ multi label/class.
Classifier

American Airlines a unit of AMR immediately matched the move, spokesman Tim Wagner said.



Feature Engineering

→ (F1)

Bag-of-words.

WM1 = {American, Airlines}

WM2 = {Tim, Wagner}

(F2)

Head words.

HM1 = Airlines

HM2 = Wagner

HM12 = Airlines + Wagner

(F3)

Left context / right context

M2 - 1 = spokesman

M2 + 1 = said

(P4)

~~Span~~ Span of bag-of-words between the two entities. → long-range ~~the~~ dependencies

{a, AMR, of, immediately, watched, move, spokesman, the, unit}

(F5) Category of Named Entity:

M1 \rightarrow ORG M2 \rightarrow PERSON }

M12 \rightarrow ORG-PERSON

(F6) Name / Nominal / Pronoun

M1: EL = Name

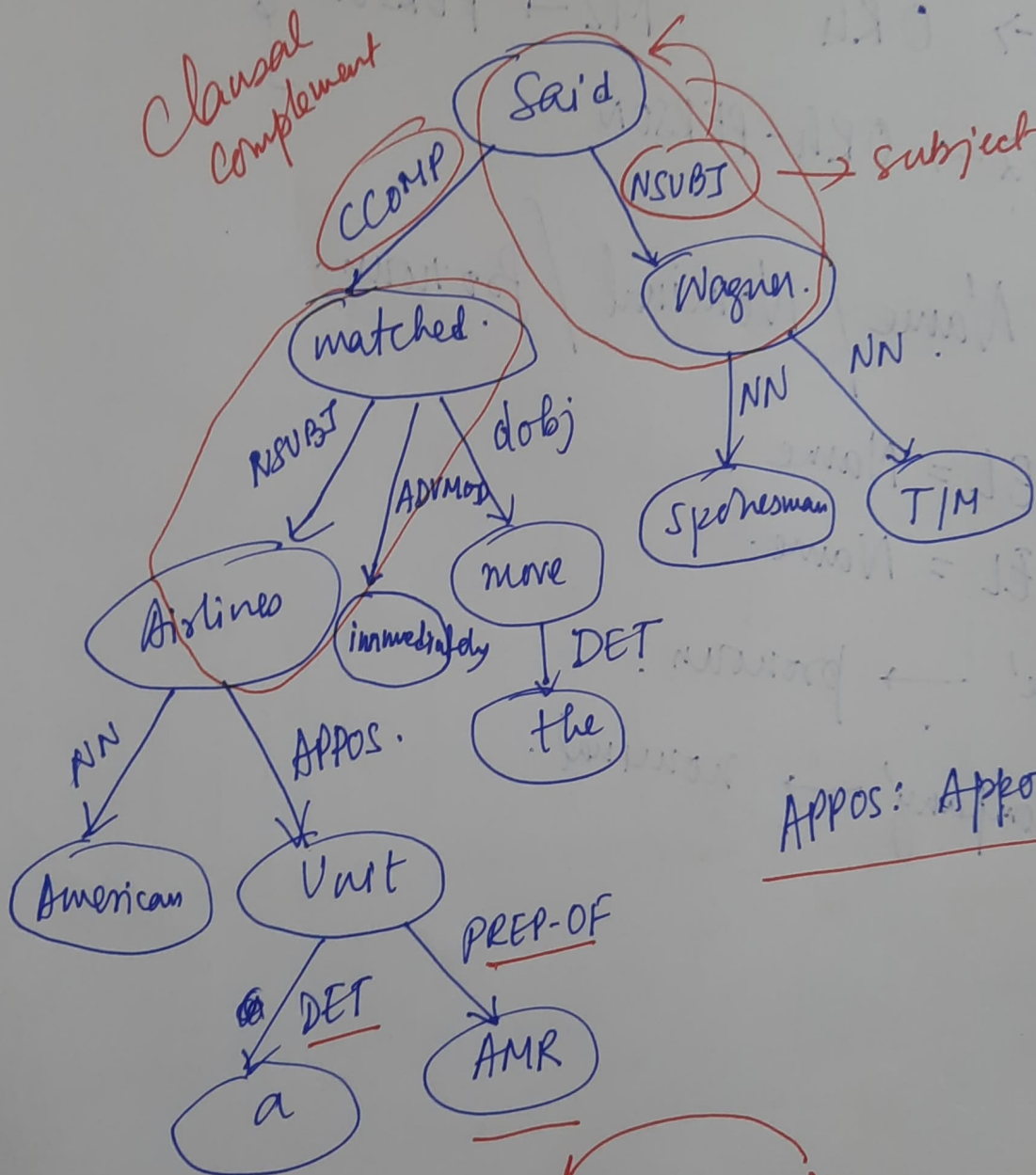
M2: EL = Name

'it', 'he' \rightarrow pronoun

'the company' \rightarrow nominal

(F7)

Dependency Parse



Appos: Appositive.

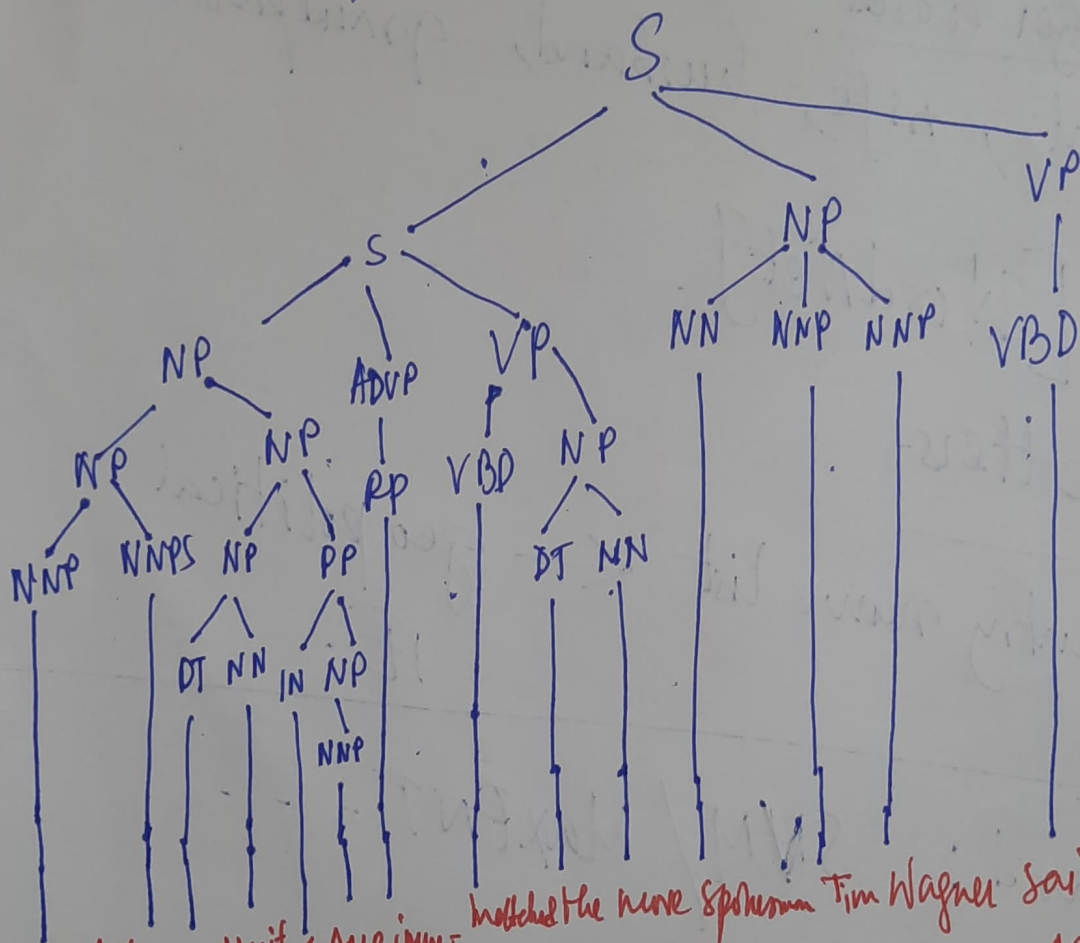
H1 ✓ DW1 = matched: Airlines.

H2 ✓ DW2 = Said: Wagner.

Path = { Airlines, matched, Said, Wagner }

F8

Constituency parse features.



watched the movie Spokenword Tim Wagner said.

NP [American Airlines] NP [a unit] PP [of] NP [AMR] ADVP [immediately]
 VP [watched] NP [the movie] NP [Spokenword] NP [Tim Wagner] VP [said.]

(79)

Trigger words:
parent, wife, husband, grandparent

[WordNet]

Gazettes

Country name list

← geopolitical

IE

Standard: SVM / MaxENT

Precision, Recall, F1-Score

Class wise precision/recall / F1

+ Macro Precision / Recall / F1