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, monastries and other important civic buildings.
X or other Y.	bruises, wounds, fractures or other injuries
Y such as X	The bow lute such as the Rambasa...
Such Y as X	such authors as Goldsmith, Wordsworth and Shakespeare.
Y including X	common-law countries including Canada and England.

$\boxed{Y}$, especially $\boxed{X}$ 	 $\boxed{\text{European countries}}$, especially $\boxed{\text{France}}$, $\boxed{\text{Germany}}$ and $\boxed{\text{Spain}}$

Microsoft Research is a part of Microsoft Corporation.

[Neither hypernymy nor hyponymy]

<u>What?</u> <u>Meronymy & holonymy</u>

basement of a tall building

part [NN-PL] of [PREP] {{the|a} [DET] [JJ|NN]}* where [NN mide. (own).

building's basement

whole [NN-PL]'s ______ part [NN-PL]

~~NN (POS)~~
JJ (POS variety).

Disadvantages of Hearst 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% (Meronymy).

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: [Mark Train, Elmira].

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

"Mark Train is buried in Elmira, NY"

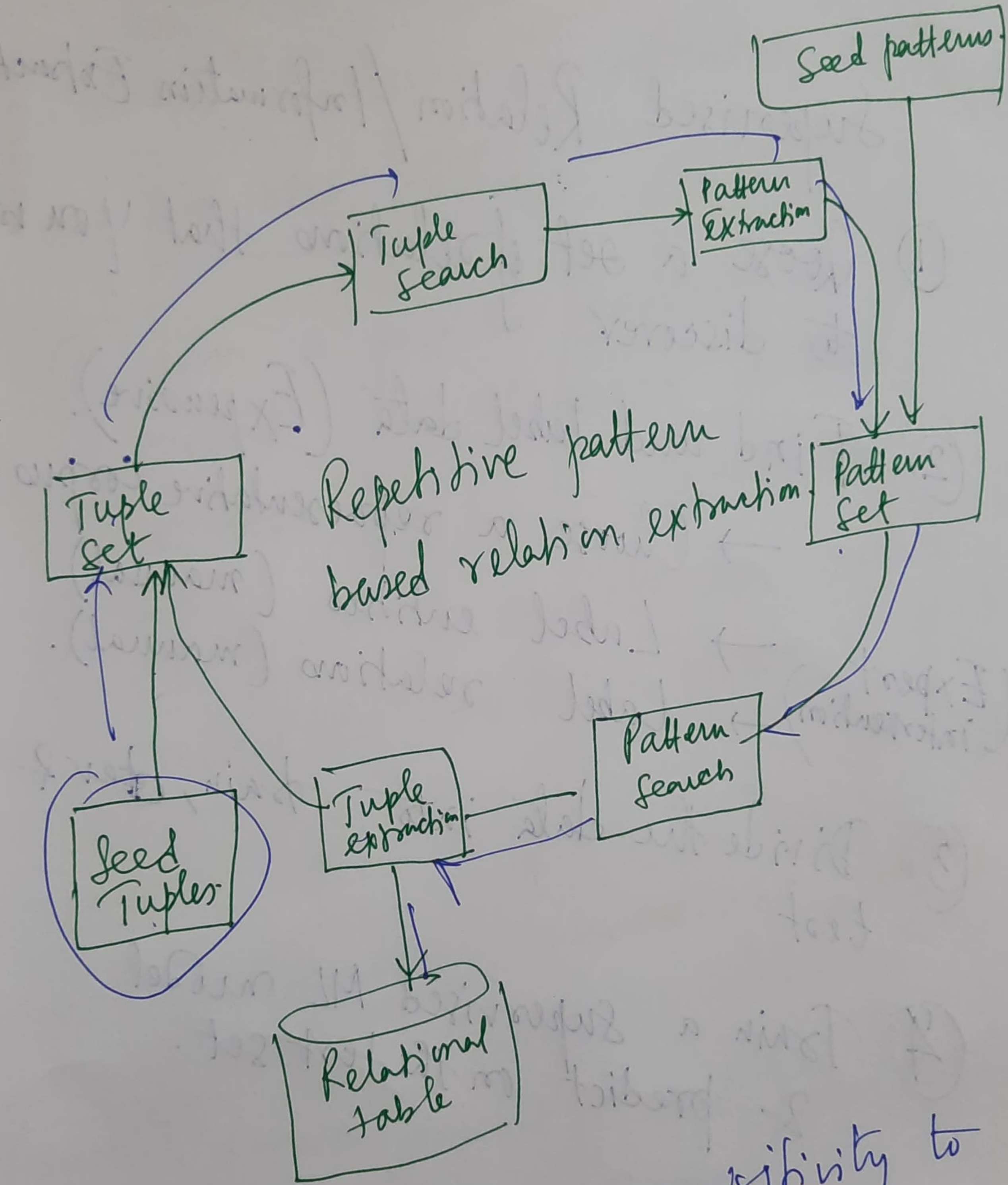
→ X is buried in Y

"The grave of Mark Train is in Elmir"

→ The grave of X is in Y

"Elmir is Mark Train's final resting place".

→ Y is X's final resting place.



Dis advantages? 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 label/class classifier

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

?

Family Citizen NIL Employment

Subsidiary FOUNDER Inventor ....

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.

F4

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

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

F5 Category of Named Entity:

M1 → ORG   M2 → PERSON }

M12 → ORG-PERSON

F6 Name / Nominal / Pronoun.

M1: EL = Name

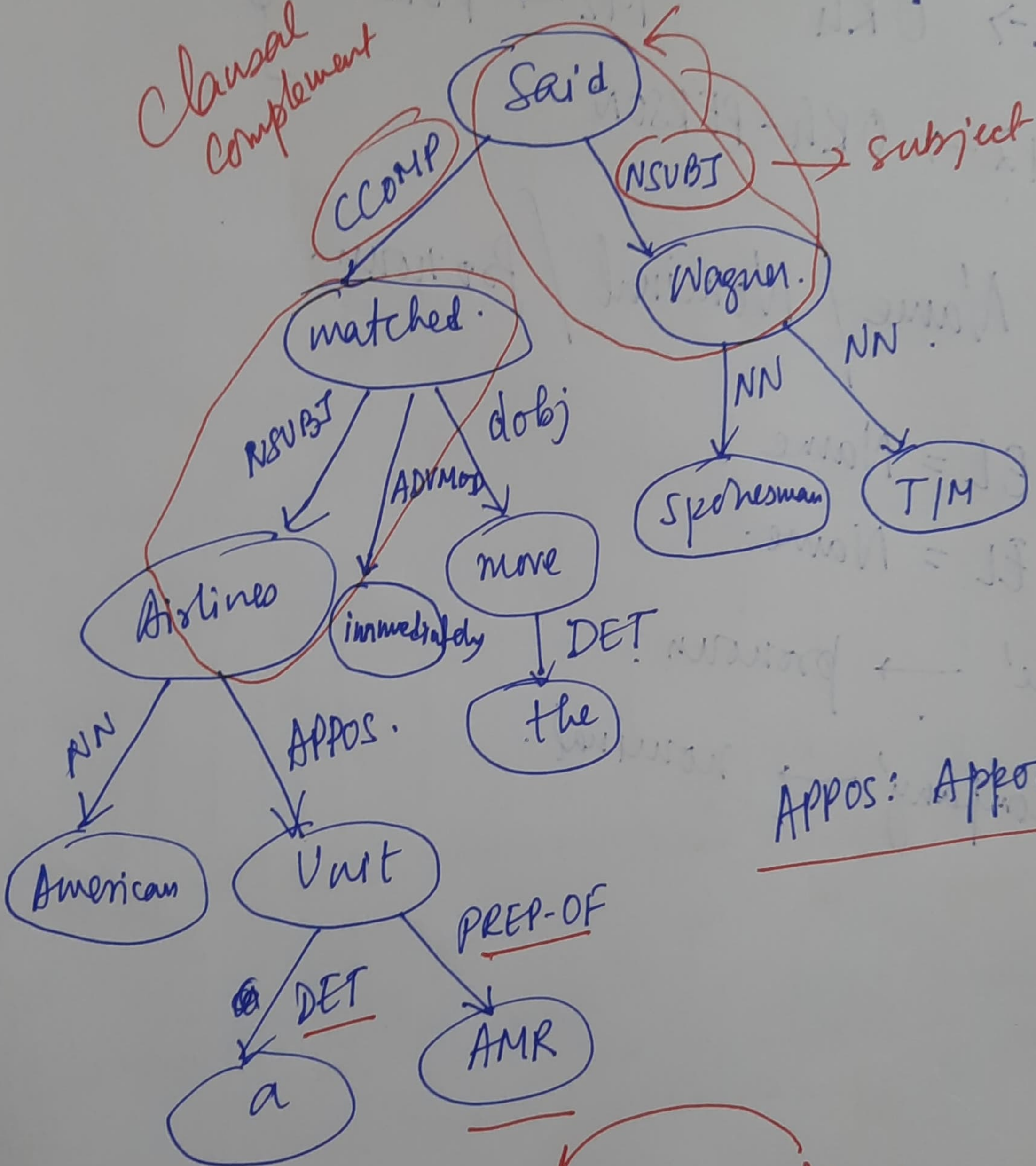
M2: EL = Name.

'It', 'he' → pronoun

'the company' → nominal.

F7

Dependency Parse

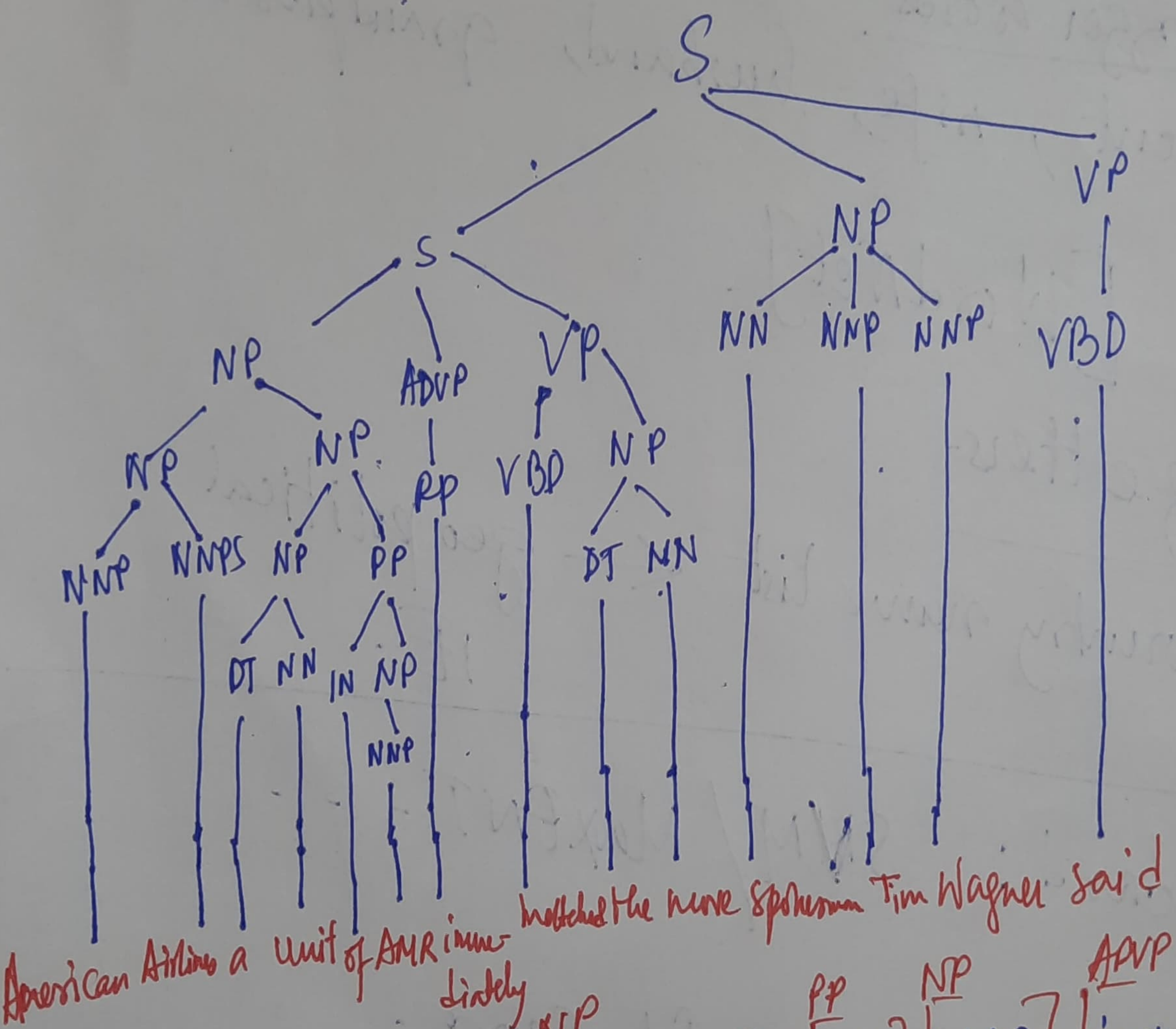


APPOS: Appositive.

H1DW1 = matched: Airlines.
H2DW2 = said: Wagner.
path = {Airlines, matched, said, Wagner}.

F8

Constituency parse features.



NP[American Airlines] NP[a unit] PP[of] NP[AMR] ADVP[immediately] VP[said]

NP[Spokesman] [Tim Wagner]

VP[matched] NP[the move]

P9

Trigger words:

parent, wife, husband, grandparent

[WordNet]

Gazetters-

Country name list ← geo political

IE

Standard: SVM / MaxENT

Precision, Recall, F1-Score.

Class wise precision/recall/
F1

+ Macro Precision/
recall / F1.