Modular approach to pattern recognition

There seem to be a few base patterns starting from a tagged term:

* nn(Noun1, term)
* amod(Adj1, term)
* appos(Noun1, term)
* appos(term, Noun1) Doesn’t actually appear in dev set, but looks plausible
* conj\_and(Noun1, term) Handles case that parse mistakes appositive for conjuction
* conj\_and(term, Noun1) ”

After getting a match for one of the base patterns, extend Noun1 to get appositive:

* appos(Noun1, Noun2)
* appos(Noun2, Noun1)
* conj\_and(Noun1, Noun2)
* conj\_and(Noun2, Noun1)

Next extend either Noun1 or Noun2 to attach “of” phrase:

* prep\_of(Noun1, Noun3)
* prep\_of(Noun2, Noun3)

Next entend the right-most noun to get parenthetical phrase. Don’t need parse for this.

* If extracted phrase so far includes “(“, extend to the next “)”
* If end of extracted phrase is followed by “(“, extend to the next “)”

Patterns in current dev set, where <term> has been tagged with a target class

<entity>, <term> appos(entity, term) Neil deGrasse Tyson, an astrophysicist

John Arterberry, executive deputy chief

Petra Koepke-Eberler, 48, a music teacher

Madame Violetta, a classical singing teacher

<term> <entity> nn(entity, term) critic Andre Bazin

President Hu

Chinese President Hu

Canadian Pacific Railway

President Bush

Persian Gulf War

<term> <entity> amod(entity, term) French critic Ander Bazin

Persian calligraphy

Iraqi government

<term> <entity> of N nn(entity, term) prep\_of(entity, N)

the Hindu festival of Holi

<entity>, <term> N amod(N, term) appos(entity, N)

Yenching University, a Christian institution

<entity>, <term> N of N2 nn(N, term) appos(N, entity) prep\_of(entity, N)

Susan Michael, the U.S. director of the

International Christian Embassey Jerusalem

<term> N (<entity>) nn(N, term) ???(N, entity)

the first Mormon president (Mitt Romney)

Patterns to recover from bad parses:

<term>, <entity> conj\_and(entity, term) community leader, Iyal al-Ashouri

[note: appos is often mis-parsed as conj\_and. This pattern may be unreliable]

<term> N, <entity> amod(N, term) conj\_and(entity, N)

Christian community leader, Iyal al-Ashouri

[note: appos is often mis-parsed as conj\_and. This pattern may be unreliable]

<entity>, <term> N nn(N, term) conj\_and(entity, N)

Daniel Peterson, a West Indian immigrant, and Olivia John

[note: another appos parsed as conj\_and]

<term> <entity> N nn(N, term) nn(N, entity)

Persian Emperor Xerxes (Brazil's Rodrigo Santoro)

<entity> -- <term> dep(term, entity) a Mormon -- Mitt Romney

Not an implicit relation – Open IE can handle it.

<entity> V in <term> handled by Open IE Peterson's mother, who also had roots in the Caribbean