Unsupervised Learning of Narrative Schemas and their Participants

Nathanael Chambers and Dan Jurafsky
Presented by Zhichao Hu
<zhu@cs.ucsc.edu>
Review: Narrative Event Chains

• Narrative (Event) Chains: partially ordered sets of events centered around a common protagonist.

• An event is a verb together with its constellation of arguments.

• An event slot is a tuple of an event and a particular argument slot (grammatical relation): <v, d>, v is a verb, d∈{subject, object, prep}
Review: Narrative Event Chains

- A chain is a tuple \((L, O)\) where \(L\) is a set of event slots and \(O\) is a partial (temporal) ordering.

\[
L = (X \text{ pleads}), (X \text{ admits}), (\text{convicted } X), (\text{sentenced } X)
\]

\[
O = \{(\text{pleads, convicted}), (\text{convicted, sentenced}), \ldots\}
\]
Review: the Case for Arguments

• The problem of Chambers & Jurafsky, 2008: chain learning and clustering is based only on the frequency with which two verbs share arguments, ignoring any features of the arguments themselves.

• Which event is most likely to occur?
  – Incorrectly favor (fly X) because it is observed during training with all five event slots.
  – Favor (charge X) because it shares many arguments with (accuse X), (search X) and (suspect X).
Review: the Case for Joint Chains

• The problem of Chambers & Jurafsky, 2008: making judgments only between protagonist arguments, one slot per event.

• Which verb is more related to arrest, convict or capture?
  – Incorrectly choose convict because of its highest score
  – Choose capture because both the objects (suspect) and subjects (police) are identical.

  police arrest suspect

  judge convict suspect

  police capture suspect
Review: Semantic Role Labeling

• Problems:
  – Most work on semantic role labeling is supervised.
  – Or require a predefined set of roles to define the domain of their probabilistic model.

• Goals:
  – Unsupervised learning.
  – Learning the roles themselves.
Typed Narrative Chains

• A Typed Narrative Chain is a partially ordered set of event slots that share an argument, but now the shared argument is a role defined by being a member of a set of types R.

• A Typed Narrative Chain: \((L, P, O)\)
  - \(L\): set of event slots
  - \(O\): partial ordering
  - \(P\): a set of argument types (head words) representing a single role

\[
L = \{(\text{hunt X}), (X \text{ use}), (\text{suspect X}), (\text{accuse X}), (\text{search X})\}
\]
\[
P = \{\text{person}, \text{government}, \text{company}, \text{criminal}, \ldots\}
\]
\[
O = \{(\text{use}, \text{hunt}), (\text{suspect}, \text{search}), (\text{suspect}, \text{accuse}) \ldots\}
\]
Argument Types

• To learn narrative chains:
  – Parse the text
  – Resolve coreference (OpenNLP)
  – Extract chains of events that share participants
    • Record counts of arguments that are observed with each pair of event slots
    • Build the referential set for each word from its coreference chain
    • represent each observed argument by the most frequent head word in its referential set
Argument Types

But for a growing proportion of U.S. workers, the troubles really set in when they apply for unemployment benefits. Many workers find their benefits challenged.
Argument Types: CorefChain by Stanford Parser

• But for a growing proportion of **U.S. workers**, the troubles really set in when **they** apply for unemployment benefits. Many **workers** find **their** benefits challenged.

• {2=CHAIN2-"a growing proportion of U.S. workers" in sentence 1],
• 4=CHAIN4-"U.S. workers" in sentence 1, "they" in sentence 1],
• 5=CHAIN5-"the troubles" in sentence 1],
• 7=CHAIN7-"unemployment benefits" in sentence 1],
• 8=CHAIN8-"Many workers" in sentence 2, "their" in sentence 2],
• 9=CHAIN9-"their benefits challenged" in sentence 2]
Review: Event Slot Similarity without Argument Types

• PMI (pointwise mutual information):

$$pmi(e(w, d), e(v, g)) = \log \frac{P(e(w, d), e(v, g))}{P(e(w, d))P(e(v, g))}$$

• Chain similarity:

$$chainsim(C, \langle f, g \rangle) = \sum_{i=1}^{n} \text{sim}(\langle e_i, d_i \rangle, \langle f, g \rangle) \quad (1)$$
Event Slot Similarity with Argument Types

- Similarity in the context of a specific argument $a$:

$$\text{sim}(\langle e, d \rangle, \langle e', d' \rangle, a) = \text{pmi}(\langle e, d \rangle, \langle e', d' \rangle) + \lambda \log \text{freq}(\langle e, d \rangle, \langle e', d' \rangle, a)$$

(2)

- Score the entire chain for a particular argument:

$$\text{score}(C, a) = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \text{sim}(\langle e_i, d_i \rangle, \langle e_j, d_j \rangle, a)$$

(3)

- Chain similarity:

$$\text{chainsim}'(C, \langle f, g \rangle) = \max_a (\text{score}(C, a) + \sum_{i=1}^{n} \text{sim}(\langle e_i, d_i \rangle, \langle f, g \rangle, a))$$

(4)
Narrative Schema: Multiple Chains
The Model

• A Narrative Schema is a set of typed narrative chains.

• A narrative schema: \( N = (E, C) \)
  – \( E \): a set of events (verbs)
    • An event:
      – verb \( v \)
      – grammatical argument positions \( D_v \subseteq \{\text{subject, object, prep}\} \)
  – \( C \): a set of typed chains over the event slots
Narrative Schema: Multiple Chains

The Model

Typed chains
Narrative Schema: Multiple Chains

Learning Narrative Schemas

- A verb is added to a schema if both its subject and object are assigned to chains in the schema with high confidence.

---

Diagram:

- *police pull_over cars* is in the schema.
- *police search defendant* is in the schema.
- *raid* is in the schema.
- *arrest* is in the schema.
- *charge* is in the schema.
- *seize* is in the schema.
- *confiscate* is in the schema.
- *detain* is in the schema.
- *deport* is in the schema.

- *defendant, nichols, smith, simpson* is not in the schema.
- *police, agent, authorities, government* is not in the schema.
- *license* is not in the schema.
- *immigrant, reporter, cavalo, migrant, alien* is not in the schema.
Narrative Schema: Multiple Chains
Learning Narrative Schemas

- Event relatedness function: is $v$ the best?

$$narsim(N, v) = \sum_{d \in D_v} \max(\beta, \max_{c \in C_N} \text{chainsim}'(c, \langle v, d \rangle))$$ (5)

For all grammatical argument positions: subject, object and prep
Maximizing among all chains

Base score
grammatical argument position
verb

Entire narrative schema
verb

$\text{chainsim}'(c, \langle v, d \rangle)$
Narrative Schema: Multiple Chains Building Schemas

• Add the best event slot based on:

  • Old:

    \[ \max_{j:0<j<m} \text{chainsim}(c, \langle v_j, g_j \rangle) \]  \hspace{1cm} (6)

  • New:

    \[ \max_{j:0<j<|v|} \text{narsim}(N, v_j) \]  \hspace{1cm} (7)

• Verbs are incrementally added to a narrative schema by strength of similarity.
### Sample Narrative Schemas

| A produce B | A ∈ \{company, inc, corp, microsoft, iraq, co, unit, maker, ...\} | B trade C | A ∈ {} |
| A sell B    | A ∈ \{drug, product, system, test, software, funds, movie, ...\} | B fall C  | B ∈ \{dollar, share, index, mark, currency, stock, yield, price, pound, ...\} |
| A manufacture B | B ∈ \{wash, heat, thinly, onion, note\} | A *quote B | A ∈ \{police, agent, officer, authorities, troops, official, investigator, ...\} |
| A *market B | A ∈ \{potato, onion, mushroom, clove, orange, gnocchi\} | B *fall C | A confiscate B | B ∈ \{suspect, government, journalist, monday, member, citizen, client, ...\} |
| A distribute B | A ∈ \{court, judge, justice, panel, osteen, circuit, nicolau, sporkin, majority, ...\} | B *rise C | A seize B |
| A -develop B | A ∈ \{law, ban, rule, constitutionality, conviction, ruling, lawmaker, tax, ...\} | B *raid B | A raid B |
| A boil B    | A ∈ \{wash, heat, thinly, onion, note\} | A search B | A search B |
| A slice B   | A ∈ \{wash, heat, thinly, onion, note\} | A arrest B | A arrest B |
| A -peel B   | B ∈ \{potato, onion, mushroom, clove, orange, gnocchi\} | A enforce B | B ∈ \{law, ban, rule, constitutionality, conviction, ruling, lawmaker, tax, ...\} |
| A saute B   | A ∈ \{court, judge, justice, panel, osteen, circuit, nicolau, sporkin, majority, ...\} | A *overturn B | A own B |
| A cook B    | A ∈ \{wash, heat, thinly, onion, note\} | A *strike_down B | A ∈ \{company, investor, trader, corp, enron, inc, government, bank, itt, ...\} |
| A chop B    | A ∈ \{wash, heat, thinly, onion, note\} | A *challenge B | A *borrow B |

Figure 5: Six of the top 20 scored Narrative Schemas. Events and arguments in italics were marked misaligned by FrameNet definitions. * indicates verbs not in FrameNet. - indicates verb senses not in FrameNet.
Evaluation: FrameNet

• FrameNet: a database of frames, structures that characterize particular situations.
  – A set of events (the verbs and nouns that describe them)
  – A set of frame-specific semantic roles called frame elements that can be arguments of the lexical units in the frame

• Qualitative evaluations using FrameNet:
  – Verb groupings
  – Linking structure
  – Argument roles
"Manufacturing" in FrameNet

FEs:

Core:

Factory [Fac] identifies the particular plant where the Product is manufactured.

Producer [Man]
Semantic Type: Sentient

Product [Pro]

Lexical Units:

assembly_line.n, fabricate.v, fabrication.n, industrial.a, make.v, maker.n, manufacture.n, manufacture.v, manufacturer.n, manufacturing.n, produce.v, producer.n, product.n, production.n
Evaluation: Cloze

• The cloze task: remove a random word from a sentence and ask the subject to guess what is missing.
• The narrative cloze is a variation on this idea that removes an event slot from a known narrative chain. Performance is measured by the position of the missing event slot in a system’s ranked guess list.
• Training and Test Data: NYT portion of the Gigaword Corpus (years 1994-2004), one million articles.
Evaluation: Cloze

Narrative Cloze Test

![Graph showing the evaluation of Narrative Cloze Test with different methods over time. The x-axis represents training data from 1994 to X, and the y-axis represents ranked position. The graph compares Chain, Typed Chain, Schema, and Typed Schema methods.]
Thank you!