Semantic Roles & Semantic Role Labeling

Ling571
Deep Processing Techniques for NLP
February 17, 2016

Roadmap

- Semantic role labeling (SRL):
  - Motivation:
    - Between deep semantics and slot-filling
  - Thematic roles
  - Thematic role resources
    - PropBank, FrameNet
  - Automatic SRL approaches

Semantic Analysis

- Two extremes:
  - Full, deep compositional semantics
    - Creates full logical form
    - Links sentence meaning representation to logical world model representation
    - Powerful, expressive, AI-complete
  - Domain-specific slot-filling:
    - Common in dialog systems, IE tasks
    - Narrowly targeted to domain/task
    - Often pattern-matching
    - Low cost, but lacks generality, richness, etc

Semantic Role Labeling

- Typically want to know:
  - *Who* did *what* to *whom*, *where*, *when*, and *how*
- Intermediate level:
  - Shallower than full deep composition
  - Abstracts away (somewhat) from surface form
  - Captures general predicate-argument structure info
  - Balance generality and specificity

Example

- Yesterday Tom chased Jerry.
- Yesterday Jerry was chased by Tom.
- Tom chased Jerry yesterday.
- Jerry was chased yesterday by Tom.

- Semantic roles:
  - Chaser: Tom
  - ChasedThing: Jerry
  - TimeOfChasing: yesterday
- Same across all sentence forms

Full Event Semantics

- Neo-Davidsonian style:
  - exists e. Chasing(e) & Chaser(e,Tom) & ChasedThing(e,Jerry) & TimeOfChasing(e,Yesterday)
- Same across all examples
- Roles: Chaser, ChasedThing, TimeOfChasing
  - Specific to verb "chase"
  - Aka "Deep roles"

Issues

- Challenges:
  - How many roles for a language?
    - Arbitrarily many deep roles
      - Specific to each verb's event structure
  - How can we acquire these roles?
    - Manual construction?
    - Some progress on automatic learning
      - Still only successful on limited domains (ATIS, geography)
  - Can we capture generalities across verbs/events?
    - Not really, each event/role is specific
- Alternative: thematic roles

Thematic Roles

- Describe semantic roles of verbal arguments
  - Capture commonality across verbs
  - E.g. subject of break, open is AGENT
    - AGENT: volitional cause
    - THEME: things affected by action
  - Enables generalization over surface order of arguments
    - $\text{John}_{\text{AGENT}}$ broke the $\text{window}_{\text{THEME}}$
    - The $\text{rock}_{\text{INSTRUMENT}}$ broke the $\text{window}_{\text{THEME}}$
    - The $\text{window}_{\text{THEME}}$ was broken by $\text{John}_{\text{AGENT}}$

Thematic Roles

- Thematic grid, θ-grid, case frame
  - Set of thematic role arguments of verb
    - E.g. Subject: AGENT; Object: THEME, or
    - Subject: INSTR; Object: THEME

- Verb/Diathesis Alternations
  - Verbs allow different surface realizations of roles
    - $\text{Doris}_{\text{AGENT}}$ gave the $\text{book}_{\text{THEME}}$ to $\text{Cary}_{\text{GOAL}}$
    - $\text{Doris}_{\text{AGENT}}$ gave $\text{Cary}_{\text{GOAL}}$ the $\text{book}_{\text{THEME}}$
  - Group verbs into classes based on shared patterns

Canonical Roles

Thematic Role	Example
AGENT	*The waiter* spilled the soup.
EXPERIENCER	*John* has a headache.
FORCE	*The wind* blows debris from the mall into our yards.
THEME	Only after Benjamin Franklin broke *the ice*...
RESULT	The French government has built a *regulation-size baseball diamond*...
CONTENT	Mona asked *"You met Mary Ann at a supermarket?"*
INSTRUMENT	He turned to poaching catfish, stunning them *with a shocking device*...
BENEFICIARY	Whenever Ann Callahan makes hotel reservations *for her boss*...
SOURCE	I flew in *from Boston*.
GOAL	I drove *to Portland*.

Thematic Role Issues

- Hard to produce
  - Standard set of roles
    - Fragmentation: Often need to make more specific
      - E,g, INSTRUMENTS can be subject or not

  - Standard definition of roles
    - Most AGENTs: animate, volitional, sentient, causal
    - But not all....

- Strategies:
  - Generalized semantic roles: PROTO-AGENT/PROTO-PATIENT
    - Defined heuristically: PropBank
  - Define roles specific to verbs/nouns: FrameNet

PropBank

- Sentences annotated with semantic roles
  - Penn and Chinese Treebank
  - Roles specific to verb sense
    - Numbered: Arg0, Arg1, Arg2,...
      - Arg0: PROTO-AGENT; Arg1: PROTO-PATIENT, etc
      - > 1: Verb-specific
  - E.g. agree.01
    - Arg0: Agreer
    - Arg1: Proposition
    - Arg2: Other entity agreeing
    - Ex1: [$_{\text{Arg0}}$The group] agreed [$_{\text{Arg1}}$it wouldn't make an offer]

Propbank

- Resources:
  - Annotated sentences
    - Started w/Penn Treebank
    - Now: Google answerbank, SMS, webtext, etc
      - Also English and Arabic
  - Framesets:
    - Per-sense inventories of roles, examples
    - Span verbs, adjectives, nouns (e.g. event nouns)
- http://verbs.colorado.edu/propbank
- Recent status:
  - 5940 verbs w/ 8121 framesets;
  - 1880 adjectives w/2210 framesets

FrameNet (Fillmore et al)

- Key insight:
  - Commonalities not just across diff't sentences w/*same* verb but across *different* verbs (and nouns and adjs)
- PropBank
  - [$_{Arg0}$Big Fruit Co.] increased [$_{Arg1}$ the price of bananas].
  - [$_{Arg1}$The price of bananas] was increased by [$_{Arg0}$ BFCo].
  - [$_{Arg1}$The price of bananas] increased [$_{Arg2}$ 5%].
- FrameNet
  - [$_{ATTRIBUTE}$The price] of [$_{ITEM}$bananas] increased [$_{DIFF}$5%].
  - [$_{ATTRIBUTE}$The price] of [$_{ITEM}$bananas] rose [$_{DIFF}$5%].
  - There has been a [$_{DIFF}$5%] rise in [$_{ATTRIBUTE}$ the price] of [$_{ITEM}$ bananas].

FrameNet

- Semantic roles specific to Frame
  - Frame: script-like structure, roles (frame elements)

  - E.g. change_position_on_scale: increase, rise
    - Attribute, Initial_value, Final_value

  - Core, non-core roles

  - Relationships b/t frames, frame elements
    - Add causative: cause_change_position_on_scale

Change of position on scale

VERBS:	dwindle	move	soar	escalation	shift
advance	edge	mushroom	swell	explosion	tumble
climb	explode	plummet	swing	fall	
decline	fall	reach	triple	fluctuation	**ADVERBS:**
decrease	fluctuate	rise	tumble	gain	increasingly
diminish	gain	rocket		growth	
dip	grow	shift	**NOUNS:**	hike	
double	increase	skyrocket	decline	increase	
drop	jump	slide	decrease	rise	

Core Roles	
ATTRIBUTE	The ATTRIBUTE is a scalar property that the ITEM possesses.
DIFFERENCE	The distance by which an ITEM changes its position on the scale.
FINAL_STATE	A description that presents the ITEM's state after the change in the ATTRIBUTE's value as an independent predication.
FINAL_VALUE	The position on the scale where the ITEM ends up.
INITIAL_STATE	A description that presents the ITEM's state before the change in the ATTRIBUTE's value as an independent predication.
INITIAL_VALUE	The initial position on the scale from which the ITEM moves away.
ITEM	The entity that has a position on the scale.
VALUE_RANGE	A portion of the scale, typically identified by its end points, along which the values of the ATTRIBUTE fluctuate.
Some Non-Core Roles	
DURATION	The length of time over which the change takes place.
SPEED	The rate of change of the VALUE.
GROUP	The GROUP in which an ITEM changes the value of an ATTRIBUTE in a specified way.

FrameNet

- Current status:
  - 1216 frames
  - ~13500 lexical units (mostly verbs, nouns)
  - Annotations over:
    - Newswire (WSJ, AQUAINT)
    - American National Corpus
- Under active development
- Still only ~6K verbs, limited coverage

Semantic Role Labeling

- Aka Thematic role labeling, shallow semantic parsing
- Form of predicate-argument extraction
- Task:
  - For each predicate in a sentence:
    - Identify which constituents are arguments of the predicate
    - Determine correct role for each argument
- Both PropBank, FrameNet used as targets
- Potentially useful for many NLU tasks:
  - Demonstrated usefulness in Q&A, IE

SRL in QA

- Intuition:
  - Surface forms obscure Q&A patterns
  - *Q: What year did the U.S. buy Alaska?*
  - S_A*:…before Russia sold Alaska to the United States in 1867*

  - Learn surface text patterns?
    - Long distance relations, require huge # of patterns to find
  - Learn syntactic patterns?
    - Different lexical choice, different dependency structure

Semantic Roles & QA

- Approach:
  - Perform semantic role labeling
    - FrameNet
  - Perform structural and semantic role matching
  - Use role matching to select answer

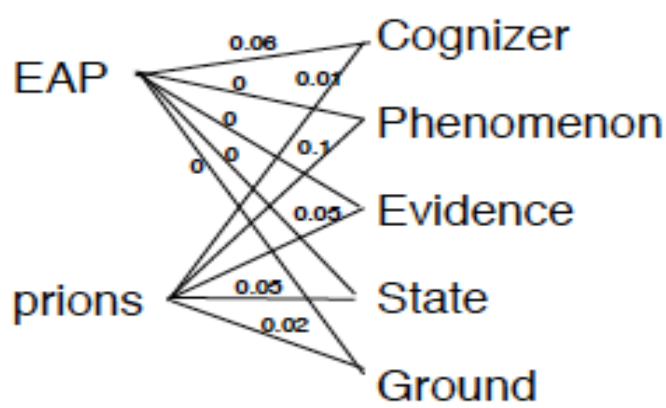
Q: Who discovered prions?

S: 1997: Stanley B. Prusiner, United States, discovery of prions, ...

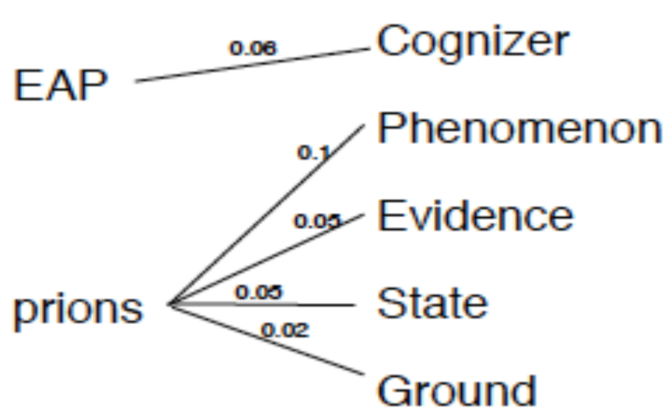
SemStruc ^{q}

p: discover

Original SR assignments:



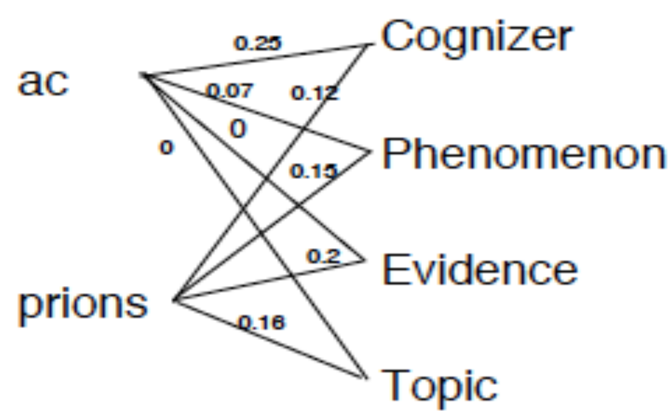
Optimized SR assignments:



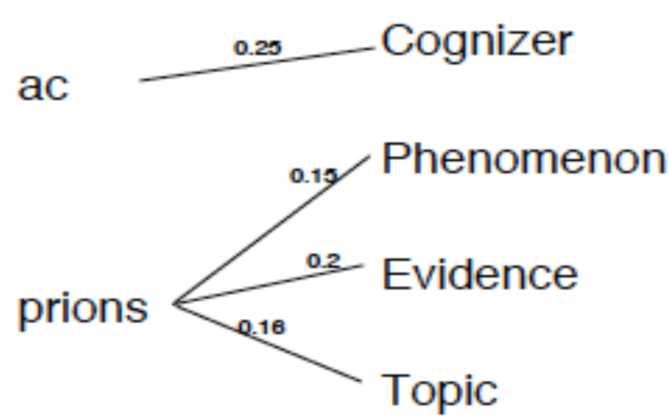
SemStruc ^{ac} (ac: Stanley B. Prusiner)

p: discovery

Original SR assignments:



Optimized SR assignments:



Summary

- FrameNet and QA:
  - FrameNet still limited (coverage/annotations)
  - Bigger problem is lack of alignment b/t Q & A frames
- Even if limited,
  - Substantially improves where applicable
  - Useful in conjunction with other QA strategies
  - Soft role assignment, matching key to effectiveness

SRL Subtasks

- Argument identification:
  - The [San Francisco Examiner] issued [a special edition] [yesterday].
  - Which spans are arguments?
    - In general (96%), arguments are (gold) parse constituents
    - 90% arguments are aligned w/auto parse constituents
- Role labeling:
  - The [$_{Arg0}$San Francisco Examiner] issued [$_{Arg1}$a special edition] [$_{ArgM\text{-}TMP}$yesterday].

Semantic Role Complexities

- Discontinuous arguments:
  - [$_{Arg1}$The pearls], [$_{Arg0}$ she] said, [$_{C\text{-}Arg1}$ are fake].
- Arguments can include referents/pronouns:
  - [$_{Arg0}$The pearls], [$_{R\text{-}Arg0}$ that] are [$_{Arg1}$ fake]

SRL over Parse Tree

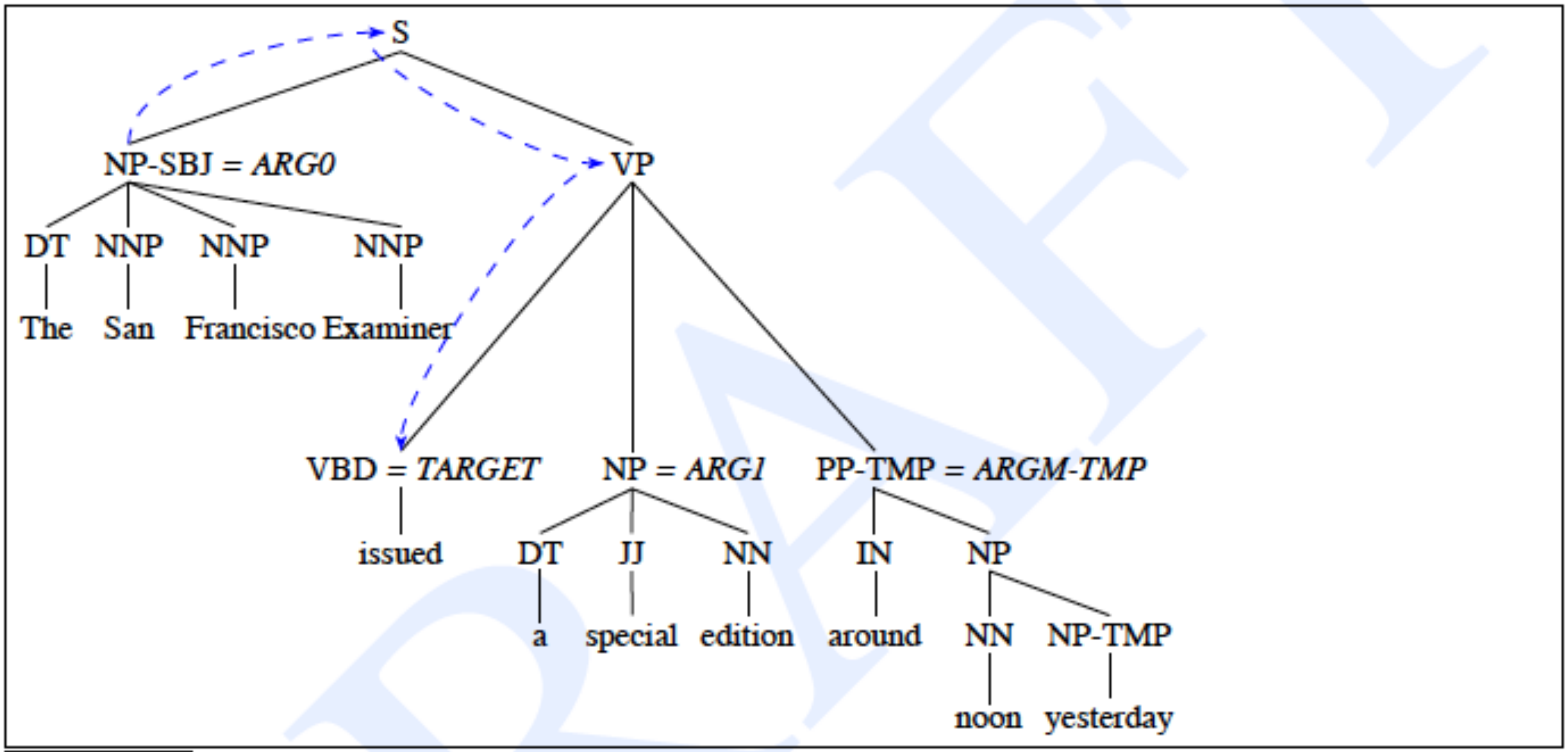


Figure 20.16 Parse tree for a PropBank sentence, showing the PropBank argument labels. The dotted line shows the path feature NP↑S↓VP↓VBD for ARG0, the NP-SBJ constituent *the San Francisco Examiner*.

Basic SRL Approach

- Generally exploit supervised machine learning
- Parse sentence (dependency/constituent)
  - For each predicate in parse:
    - For each node in parse:
      - Create a feature vector representation
      - Classify node as semantic role (or none)
- Much design in terms of features for classification

Classification Features

- Gildea & Jurafsky, 2002 (foundational work)
  - Employed in most SRL systems
- Features:
  - specific to candidate constituent argument
  - for predicate generally
- Governing **predicate**:
  - Nearest governing predicate to the current node
    - Verbs usually (also adj, noun in FrameNet)
    - E.g. ‘issued’
  - Crucial: roles determined by predicate

SRL Features

- Constituent internal information:
  - Phrase type:
    - Parse node dominating this constituent
      - E.g. NP
    - Different roles tend to surface as different phrase types
  - Head word:
      - E.g. Examiner
      - Words associated w/specific roles – e.g. pronouns as agents
  - POS of head word:
      - E.g. NNP

SRL Features

- Structural features:
  - Path: Sequence of parse nodes from const to pred
    - E.g. $\mathrm{NP}\uparrow\mathrm{S}\downarrow\mathrm{VP}\downarrow\mathrm{VBD}$
      - Arrows indicate direction of traversal
    - Can capture grammatical relations
  - Linear position:
    - Binary: Is constituent **before** or **after** predicate
      - E.g. before
  - Voice:
    - Active or passive of clause where constituent appears
      - E.g. active (strongly influences other order, paths, etc)
  - Verb subcategorization

Other SRL Constraints

- Many other features employed in SRL
  - E.g. NER on constituents, neighboring words, path info
- Global Labeling constraints:
  - Non-overlapping arguments:
    - FrameNet, PropBank both require
  - No duplicate roles:
    - Labeling of constituents is not independent
      - Assignment to one constituent changes probabilities for others

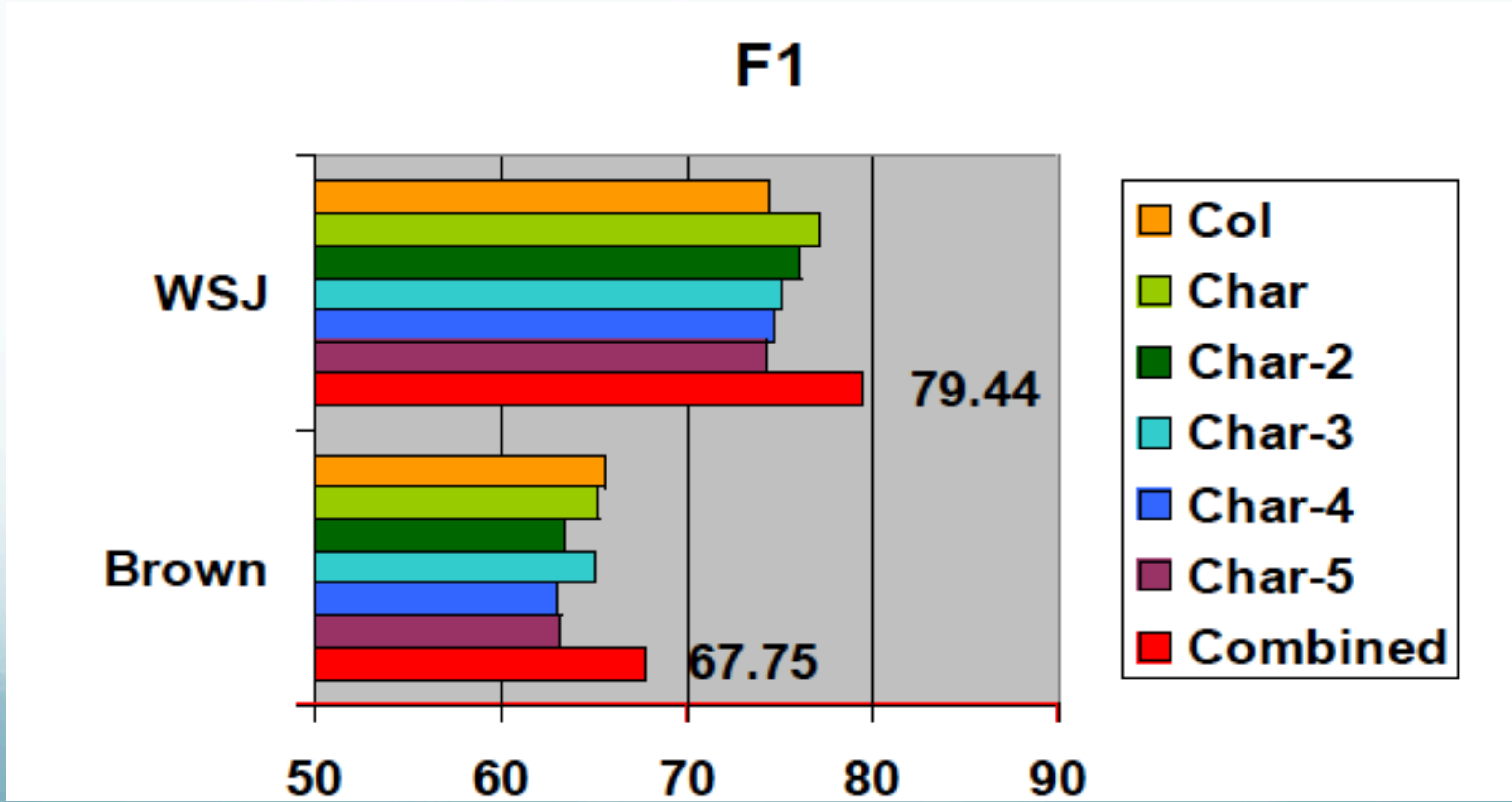
Classification Approaches

- Many SRL systems use standard classifiers
  - E.g. MaxEnt, SVM
  - However, hard to effectively exploit global constraints

- Alternative approaches
  - Classification + reranking
  - Joint modeling
  - Integer Linear Programming (ILP)
    - Allows implementation of global constraints over system

State-of-the-Art

- Best system from CoNLL shared task (PropBank)
  - ILP-based system (Punyakanok)



FrameNet “Parsing”

- (Das et al., 2014)
- Identify targets that evoke frames
  - ~ 79.2% F-measure
- Classify targets into frames
  - 61% for exact match
- Identify arguments
  - ~ 50%

SRL Challenges

- Open issues:
  - SRL degrades significantly across domains
    - E.g. WSJ → Brown: Drops > 12% F-measure
  - SRL depends heavily on effectiveness of other NLP
    - E.g. POS tagging, parsing, etc
    - Errors can accumulate
  - Coverage/generalization remains challenging
    - Resource coverage still gappy (FrameNet, PropBank)
- Publicly available implementations:
  - Shalmaneser, SEMAFOR