School of CIT -
Social Computing
Research Group

Technical
University
of Munich

Natural Language Processing
IN2361

Prof. Dr. Georg Groh

Chapter 19
Semantic Role Labeling

- content is based on [1]
- certain elements (e.g. equations or tables) were taken over or taken over in a modified form from [1]
- citations of [1] or from [1] are omitted for legibility
- errors are fully in the responsibility of Georg Groh
- BIG thanks to Dan and James for a great book!

Semantic Role Labeling

- XYZ corporation bought the stock.
- They sold the stock to XYZ corporation.
- The stock was bought by XYZ corporation.
- The purchase of the stock by XYZ corporation...
- The stock purchase by XYZ corporation...

∃purchase event; participants: XYZ Corp, some stock;
XYZ Corp: role of acquiring the stock

- Semantic roles: representations expressing abstract roles of arguments of a predicate of an event: semantic properties + relation to syntactic role in sentence

- Semantic role labeling: task of assigning roles to the constituents or phrases in sentences.

- Selectional restrictions: semantic sortal restrictions or preferences an individual predicate can express about its potential arguments, e.g. theme of the verb eat is generally something edible

Semantic Roles

- logic-based semantics representation of sentences:

Sasha broke the window.
Pat opened the door.

$\Rightarrow$

$$\exists e,x,y\ Breaking(e) \wedge Breaker(e, Sasha) \wedge BrokenThing(e,y) \wedge Window(y)$$
$$\exists e,x,y\ Opening(e) \wedge Opener(e, Pat) \wedge OpenedThing(e,y) \wedge Door(y)$$

- → Deep roles: of subject of verb brake: Breaker;
  of subject of verb brake: BrokenThing;
  deep roles are specific to Breaking event

- → more abstract Thematic roles: Breaker, Opener → Agents;
  BrokenThing, OpenedThing → Themes

Thematic Role	Definition
AGENT	The volitional causer of an event
EXPERIENCER	The experiencer of an event
FORCE	The non-volitional causer of the event
THEME	The participant most directly affected by an event
RESULT	The end product of an event
CONTENT	The proposition or content of a propositional event
INSTRUMENT	An instrument used in an event
BENEFICIARY	The beneficiary of an event
SOURCE	The origin of the object of a transfer event
GOAL	The destination of an object of a transfer event

Semantic 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 city built a *regulation-size baseball diamond*...
CONTENT	Mona asked *"You met Mary Ann at a supermarket?"*
INSTRUMENT	He poached 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*.

- deep roles, thematic roles: semantic roles

Diathesis Alternations

- Semantic roles:
  - act as shallow meaning representation allowing simple inferences not possible from pure surface string of words
  - help generalize over different surface realizations of predicate arguments

John broke the window.
AGENT THEME

John broke the window with a rock.
AGENT THEME INSTRUMENT

The rock broke the window.
INSTRUMENT THEME

The window broke.
THEME

The window was broken by John.
THEME AGENT

→ break: (at least) possible arguments AGENT , THEME , and INSTRUMENT .

- set of thematic role arguments taken by a verb: ”thematic grid” / “θ-grid” / “case frame”

Diathesis Alternations

- some realizations of arguments of *break*:

  AGENT/Subject, THEME/Object
  AGENT/Subject, THEME/Object, INSTRUMENT/PP_{with}
  INSTRUMENT/Subject, THEME/Object
  THEME/Subject

- example realizations of *give*:

Doris	*gave*	*the book*	*to*	*Cary.*		
AGENT		THEME		GOAL		

Doris	*gave*	*Cary*	*the book.*
AGENT		GOAL	THEME

- these multiple argument structure realizations: "verb alternations" / "diathesis alternations"

Diathesis Alternations

- example for *give* above: "dative alternations" (also in "verbs of future having" (*advance, allocate, offer, owe*), "send verbs" (*forward, hand, mail*), "verbs of throwing" (*kick, pass, throw*) etc.

Doris	*gave the book*	*to Cary.*	
AGENT	THEME	GOAL	
Doris	*gave Cary*	*the book.*	
AGENT	GOAL	THEME	

- Levin (1993): semantic classes for 3100 English verbs (47 high-level classes, divided into 193 more specific classes) + their alternations + (VerbNet (2000)) links to WordNet and FrameNet entries.

Problems with Thematic Roles

- fixed set of thematic roles: sometimes too general:
  e.g. INSTRUMENT → INTERMEDIARY_INSTRUMENT, ENABLING_INSTRUMENT

  The cook opened the jar with the new gadget.
  The new gadget opened the jar.

  Shelly ate the sliced banana with a fork.
  *The fork ate the sliced banana.

  sometimes too constrained
  → other sets with many fewer or many more roles

- difficult to formally define the thematic roles:
  - definition with sufficient & necessary logical conditions: too inflexible
  - lists of necessary properties: also problematic: e.g. AGENTS are: animate, volitional, sentient, causal; but some individual agent-like NP might not exhibit any of these properties

Problems with Thematic Roles

- → variant 1: use more abstract roles: e.g. PROTO-AGENT, PROTO-PATIENT: defined by list of heuristic features that hold more for one role than for another role; the more features are fulfilled: the higher the probability to be instance of that role:
  - PROTO-AGENT: being volitionally involved in event, causing an event or a change of state in another participant, being sentient or intentionally involved, moving, etc.
  - PROTO-PATIENT: undergoing change of state, causally affected by another participant, stationary relative to other participants, etc.

- → variant 2: use roles specific to one verb or a group of semantically related verbs

- PropBank: both variants.
  FrameNet: semantic roles specific for frames

PropBank

- PropBank: role-labels all sentences in Penn TreeBank
- each verb-sense: specific set of numbered roles (frame-file): *typically*:
  - Arg0: PROTO-AGENT
  - Arg1: PROTO-PATIENT
  - Arg2: benefactive, instrument, attribute, or end state
  - Arg3: start point, benefactive, instrument, or attribute
  - Arg4: end point.

benefactive: entity affected (having an advantage or as a victim)

fall.01

Arg1:	Logical subject, patient, thing falling
Arg2:	Extent, amount fallen
Arg3:	start point
Arg4:	end point, end state of arg1
Ex1:	[$_{\text{Arg1}}$ Sales] *fell* [$_{\text{Arg4}}$ to \$25 million] [$_{\text{Arg3}}$ from \$27 million].
Ex2:	[$_{\text{Arg1}}$ The average junk bond] *fell* [$_{\text{Arg2}}$ by 4.2%].

agree.01

Arg0:	Agreer
Arg1:	Proposition
Arg2:	Other entity agreeing
Ex1:	[$_{\text{Arg0}}$ The group] *agreed* [$_{\text{Arg1}}$ it wouldn't make an offer].
Ex2:	[$_{\text{ArgM-TMP}}$ Usually] [$_{\text{Arg0}}$ John] *agrees* [$_{\text{Arg2}}$ with Mary] [$_{\text{Arg1}}$ on everything].

PropBank

- further example:

increase.01 "go up incrementally"
Arg0: causer of increase
Arg1: thing increasing
Arg2: amount increased by, EXT, or MNR
Arg3: start point
Arg4: end point

[$_{\text{Arg0}}$ Big Fruit Co.] increased [$_{\text{Arg1}}$ the price of bananas].
[$_{\text{Arg1}}$ The price of bananas] was increased again [$_{\text{Arg0}}$ by Big Fruit Co.]
[$_{\text{Arg1}}$ The price of bananas] increased [$_{\text{Arg2}}$ 5%].

- additional arguments (ArgM-TMP, ArgM-LOC etc.) (valid for most verbs):

TMP	when?	yesterday evening, now
LOC	where?	at the museum, in San Francisco
DIR	where to/from?	down, to Bangkok
MNR	how?	clearly, with much enthusiasm
PRP/CAU	why?	because ... , in response to the ruling
REC	recursive	themselves, each other
ADV	miscellaneous	
PRD	secondary predication	...ate the meat raw

PropBank

- also possible: nominal predicates: *Apple's agreement with IBM* : Arg0: Apple, Arg2: IBM → NomBank (2004)

FrameNet

- previous example with *increase*: nice; but what about

  [$_{\text{Arg1}}$ The price of bananas] increased [$_{\text{Arg2}}$ 5%].
  [$_{\text{Arg1}}$ The price of bananas] rose [$_{\text{Arg2}}$ 5%]. different verb
  There has been a [$_{\text{Arg2}}$ 5%] rise [$_{\text{Arg1}}$ in the price of bananas]. noun

- PropNet: verb-specific roles; FrameNet: Frame-specific roles

- example:
  Frame: ←→ {*reservation, flight, travel, buy, price, cost, fare, rates,plane,…*}:
  - diverse semantic relations (hyponymy, synonymy, …) between those concepts + diverse roles that those have with concrete predicates BUT
  - more important for goal above: all defined with respect to a coherent chunk of common-sense background information concerning air travel.

FrameNet

- goal: make inferences even across semantic roles of specific predicates

[$_{\text{Arg1}}$ The price of bananas] increased [$_{\text{Arg2}}$ 5%].
[$_{\text{Arg1}}$ The price of bananas] rose [$_{\text{Arg2}}$ 5%].
There has been a [$_{\text{Arg2}}$ 5%] rise [$_{\text{Arg1}}$ in the price of bananas].

- → frame: background knowledge structure that defines:
  - set of frame-specific semantic roles (frame elements)
  - set of predicates (words) that use these roles.

- relationship between frames: e.g.
  - inheritance of frame elements
  - causation: *cause_change_of position_on_a_scale* frame linked to *change_of_position_on_a_scale* frame by the *cause* relation, adding AGENT role:

[$_{\text{AGENT}}$ They] *raised* [$_{\text{ITEM}}$ the price of their soda] [$_{\text{DIFFERENCE}}$ by 2%]

FrameNet example frame: **change_position_on_a_scale**

example sentences:

[ITEM Oil] *rose* [ATTRIBUTE in price] [DIFFERENCE by 2%].

[ITEM It] has *increased* [FINAL_STATE to having them 1 day a month].

[ITEM Microsoft shares] *fell* [FINAL_VALUE to 7 5/8].

[ITEM Colon cancer incidence] *fell* [DIFFERENCE by 50%] [GROUP among men].

a steady *increase* [INITIAL_VALUE from 9.5] [FINAL_VALUE to 14.3] [ITEM dividends]

a [DIFFERENCE 5%] [ITEM dividend] *increase*...

predicates:

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 (specific to frame)	
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 (also apply to other frames (compare the ARG-M arguments of PropBank))	
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.

gloss:

This frame consists of words that indicate the change of an Item's position on a scale (the Attribute) from a starting point (Initial value) to an end point (Final value).

Semantic Role Labeling

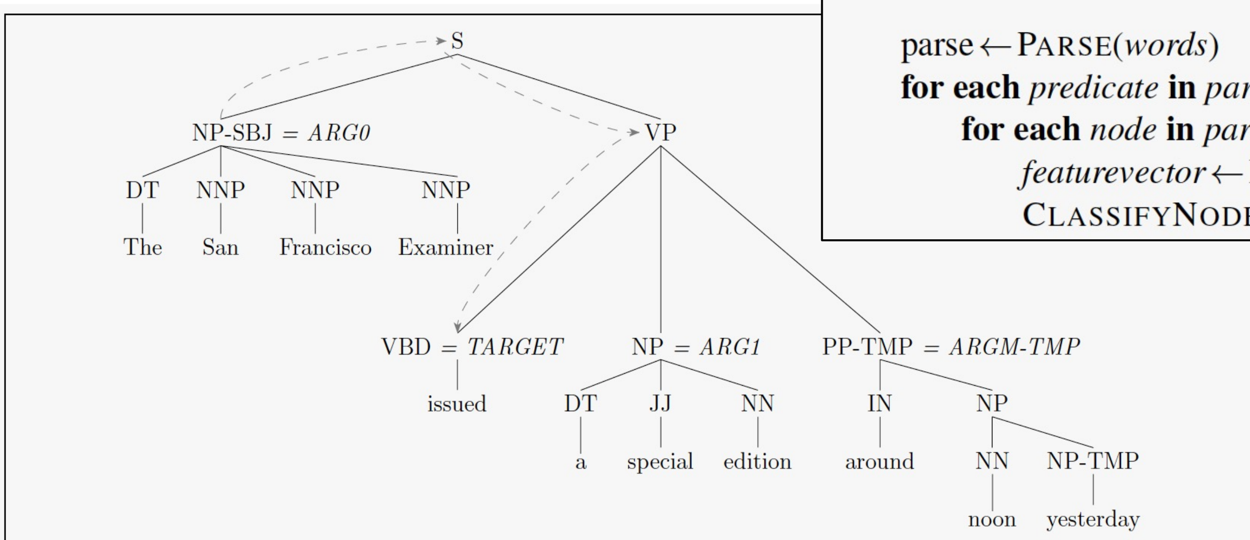
Semantic role labeling (SRL): the task of automatically finding the semantic roles of each argument of each predicate in a sentence.

→ use supervised N-class classifier, using FrameNet's or ProbBank's N-1 roles (+ "none") and e.g. restrict to classifying only their predicates

[You]	can't	[blame]	[the program]	[for being unable to identify it]
COGNIZER		TARGET	EVALUEE	REASON

[The San Francisco Examiner]	issued	[a special edition]	[yesterday]
ARG0	TARGET	ARG1	ARGM-TMP

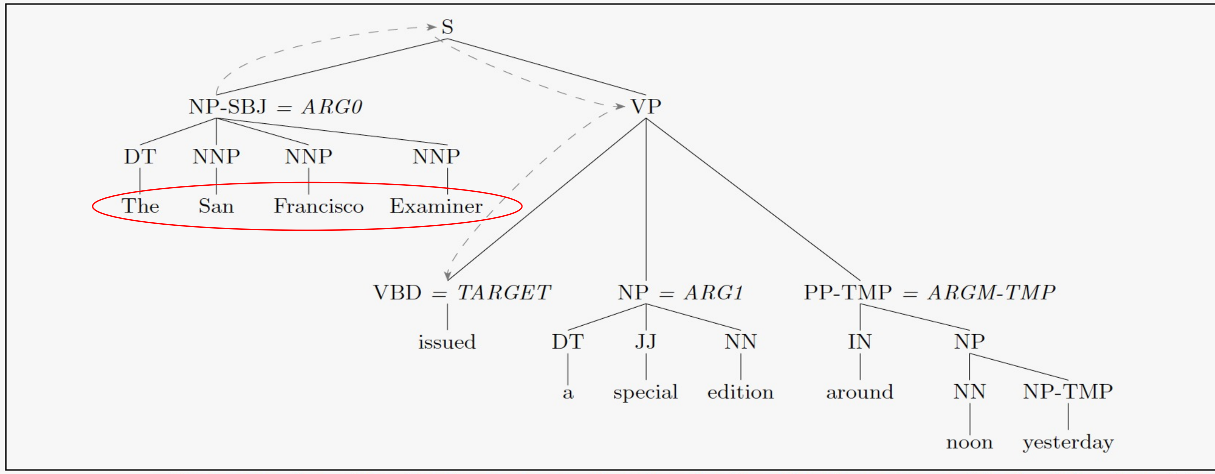
```
function SEMANTICROLELABEL(words) returns labeled tree

   parse ← PARSE(words)
   for each predicate in parse do
      for each node in parse do
         featurevector ← EXTRACTFEATURES(node, predicate, parse)
         CLASSIFYNODE(node, featurevector, parse)
```



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.*

Features for SRL

example: features of NP-SBJ constituent *The San Francisco Examiner*

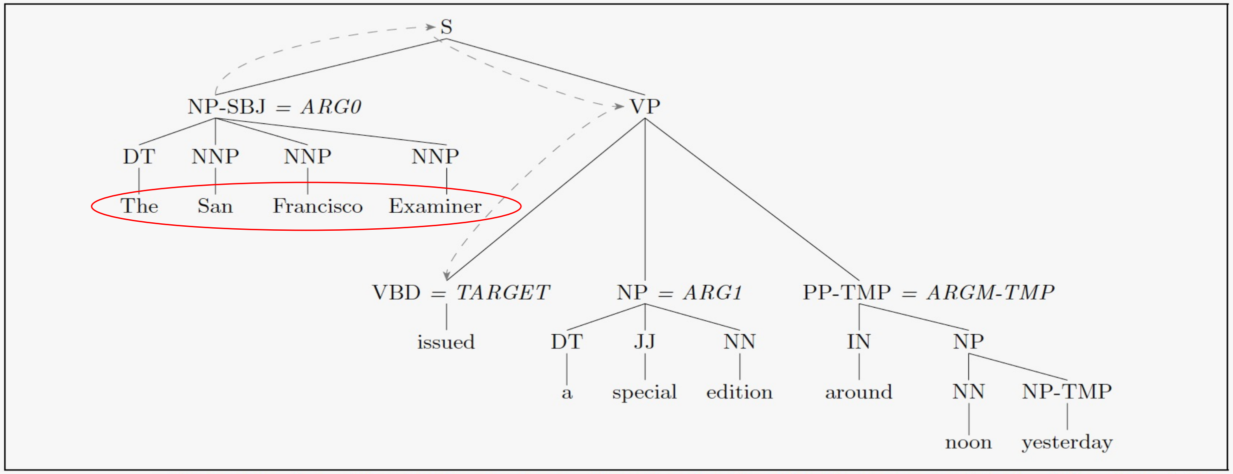


Parse tree for a PropBank sentence, showing the PropBank argument labels. The dotted line shows the **path** feature NP↑S↓VP↓VBD for ARG0, the NP-SBJ constituent *The San Francisco Examiner.*

- governing predicate (here: *issued*)
  crucial feature: labels only defined w.r.t. certain predicate
- phrase type (here: NP (or NP-SBJ)).
  Some semantic roles tend to appear as NPs, others as S or PP etc.
- headword (here: *Examiner*) (e.g. computed with head rules).
  Certain headwords (e.g., pronouns) place strong constraints on the possible semantic roles they are likely to fill.
- POS of headword (here NNP)
- path from constituent to predicate (here: NP↑S↓VP↓VBD).
  ↓ and ↑ represent upward and downward movement in the tree, respectively. path: useful as compact representation of many kinds of grammatical function relationships between constituent and predicate.

Features for SRL

example: features of NP-SBJ constituent *The San Francisco Examiner*



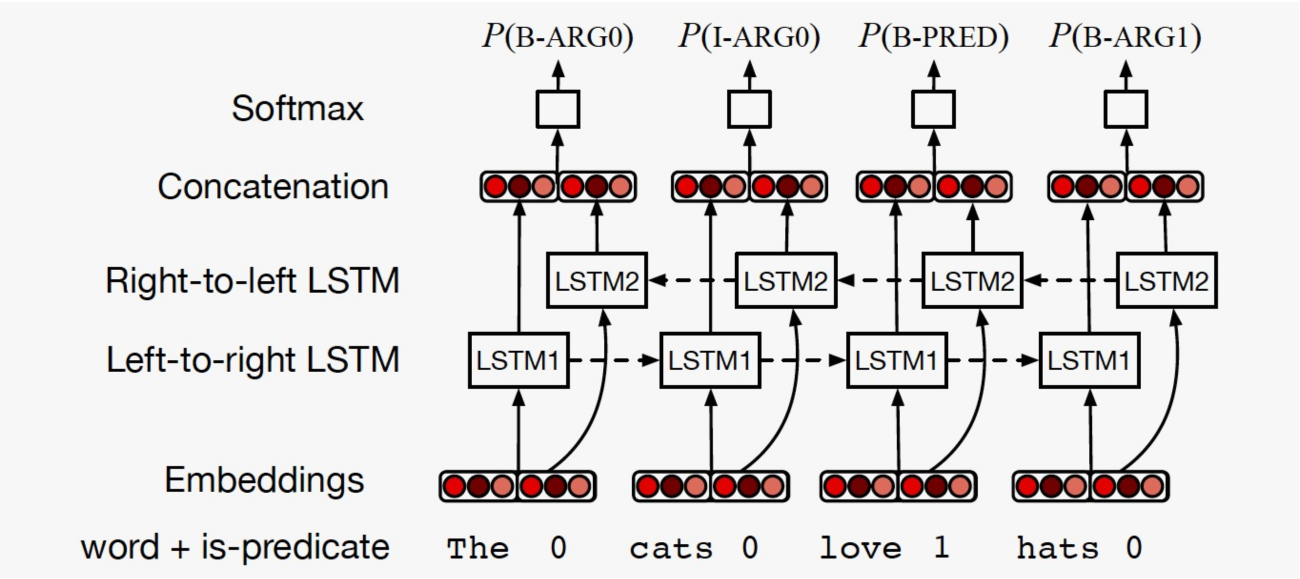
Parse tree for a PropBank sentence, showing the PropBank argument labels. The dotted line shows the **path** feature NP↑S↓VP↓VBD for ARG0, the NP-SBJ constituent *The San Francisco Examiner*.

- voice of clause in which constituent appears (here: active)
  (active or passive). Passive sentences: often strongly different linkings of semantic roles to surface form than active ones.
- position w.r.t. predicate (either before or after) (here before).
- subcategorization of predicate (here: VP → VBD NP PP)
  == the set of expected arguments that appear in the verb phrase (use phrase-structure rule that expands the immediate parent of the predicate)
- named entity type (here: ORG)
- first words and last word (here: *The, Examiner*)

Further Issues in SRL

- possible: multi-stage approach:
  - heuristics based pruning of nodes unlikely to be arguments to predicate
  - binary classifier: each node → {argument, not_argument}
  - 1-of-N classifier of arguments to roles
- pro-argument for separating argument identification from argument classification vs end-to-end: different tasks may require different features
- local classifier (simpler) vs sequence based classifier:
  - local e.g. cannot handle overlapping arguments correctly. PropBank, FrameNet: no overlapping arguments allowed
  - roles are not probabilistically independent
- add fourth stage using local probabilities to arrive at globally optimal solution (e.g. based on candidate probabilities from prev steps and Viterbi decoding etc.)

NN Approach to SRL



- sequence tagger: → $\hat{y} = \underset{y \in T}{\operatorname{argmax}} P(\mathbf{y}|\mathbf{w})$ $\quad \hat{y_i} = \underset{t \in tags}{\operatorname{argmax}} P(t|w_i)$
- but: tags are not independent (e.g. B-ARGO before I-ARG0) → use CRF instead of softmax AND/ OR
- use Viterbi decoding on the fully probability distribution for each tag, setting the transition probabilities for forbidden transitions to 0.

Modern NN Approach to SRL

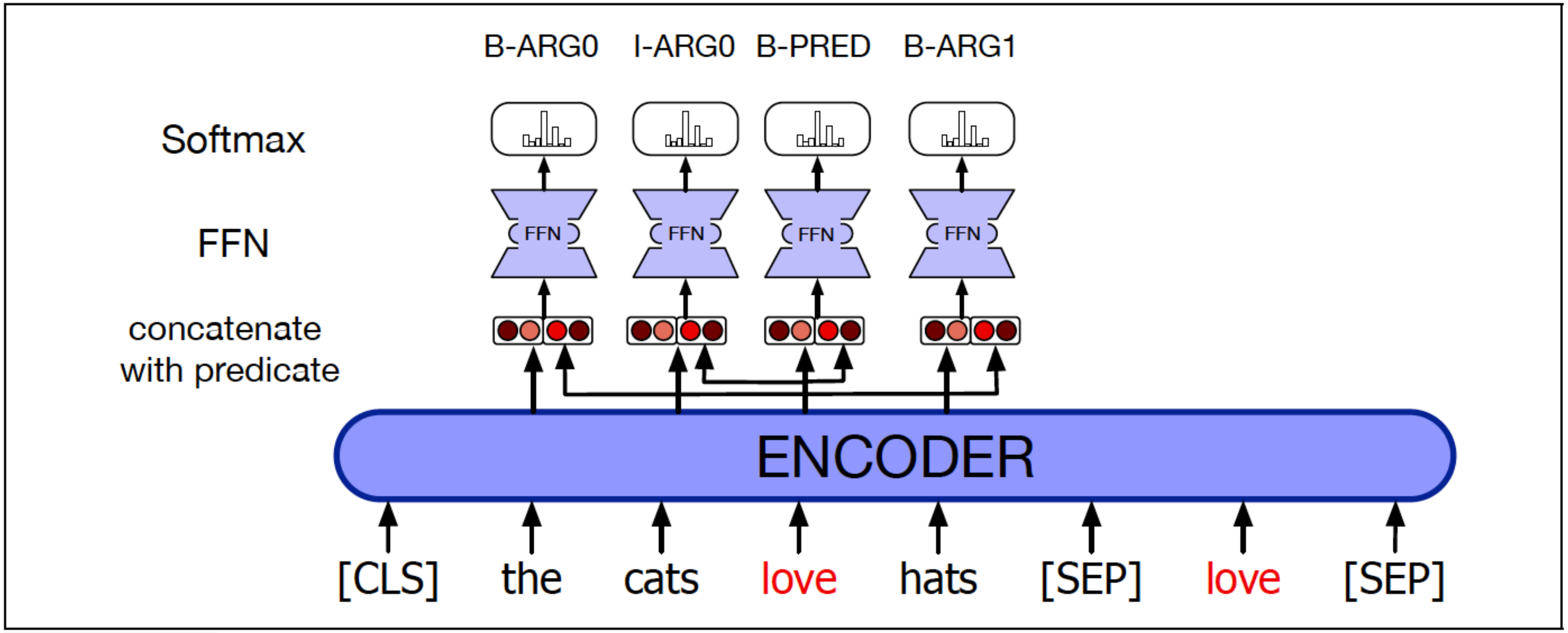


Figure 19.6 A simple neural approach to semantic role labeling. The input sentence is followed by [SEP] and an extra input for the predicate, in this case *love*. The encoder outputs are concatenated to an indicator variable which is 1 for the predicate and 0 for all other words After He et al. (2017) and Shi and Lin (2019).

Selectional Restrictions

- selectional restriction: semantic type constraint that a verb-sense imposes on the kind of concepts that are allowed to fill its argument roles.

  I want to eat someplace nearby.

  I want to eat Malaysian food.

  The restaurant serves green-lipped mussels.

  Which airlines serve Denver?

  In rehearsal, I often ask the musicians to imagine a tennis game.

  Radon is an odorless gas that can't be detected by human senses.

  To diagonalize a matrix is to find its eigenvalues.

- → many different cases

Selectional Restrictions

- first order logic: computationally too costly, not enough common-sense knowledge data-base coverage, too expressive

$$\exists e,x,y\ Eating(e) \wedge Agent(e,x) \wedge Theme(e,y)$$

$$\exists e,x,y\ Eating(e) \wedge Agent(e,x) \wedge Theme(e,y) \wedge EdibleThing(y)$$

$$\exists e,x,y\ Eating(e) \wedge Eater(e,x) \wedge Theme(e,y) \wedge EdibleThing(y) \wedge Hamburger(y)$$

- → use WordNet synsets:
  - predicate specifies WordNet synset as selectional restriction on each of its arguments.
  - meaning representation is well-formed if the role filler word is a hyponym (subordinate) of this synset.
  - example: selectional restriction on THEME role of *eat*: synset {food, nutrient}, glossed as: “any substance that can be metabolized by an animal to give energy and build tissue” .

```
Sense 1
hamburger, beefburger --
(a fried cake of minced beef served on a bun)
=> sandwich
   => snack food
      => dish
         => nutriment, nourishment, nutrition...
            => food, nutrient
               => substance
                  => matter
                     => physical entity
                        => entity
```

Selectional Restrictions and Preferences

- too many exceptions to such “rules”

  But it fell apart in 1931, perhaps because people realized you can’t eat gold for lunch if you’re hungry

  In his two championship trials, Mr. Kulkarni ate glass on an empty stomach, accompanied only by water and tea.

  → use soft constraints / preferences instead of hard constraints

- → define selectional preference strength: general amount of information that a predicate tells us about the semantic class of its arguments;
  e.g. *eat*: has high preference strength, *be*: has low preference strength

  selectional association: relative contribution of a class to the general selectional preference of the verb:

Verb	Direct Object Semantic Class	Assoc	Direct Object Semantic Class	Assoc
read	WRITING	6.80	ACTIVITY	-.20
write	WRITING	7.26	COMMERCE	0
see	ENTITY	5.79	METHOD	-0.01

Selectional Restrictions

- idea: compare: $P(c)$ (how likely is it that a direct object will fall into class c) and
  $P(c|v)$ (how likely is it that the direct object of the specific verb v
  will fall into semantic class c).

  use KL divergence for comparison:

$$D(P||Q) = \sum_x P(x) \log \frac{P(x)}{Q(x)}$$

  - Resnik (1993): sectional preference: how much information, in bits, does verb v express about the possible semantic class of its argument:

$$\begin{aligned} S_R(v) &= D(P(c|v)||P(c)) \\ &= \sum_c P(c|v) \log \frac{P(c|v)}{P(c)} \end{aligned}$$

- → selectional association: relative contribution of a class to the general selectional preference of the verb (compare PMI):

$$A_R(v,c) = \frac{1}{S_R(v)} P(c|v) \log \frac{P(c|v)}{P(c)}$$

Selectional Preference via Conditional Probability

- even simpler: use conditional probability of an argument noun given a verb for a particular relation $P(n|v,r)$ as selectional preference metric for that pair of words:

$$P(n|v,r) = \begin{cases} \frac{C(n,v,r)}{C(v,r)} & \text{if } C(n,v,r) > 0 \\ 0 & \text{otherwise} \end{cases}$$

- also possible: use:

$$P(v|n,r) = \begin{cases} \frac{C(n,v,r)}{C(n,r)} & \text{if } C(n,v,r) > 0 \\ 0 & \text{otherwise} \end{cases}$$

- also possible: if parses are not available: use POS patterns such as V Det N to approximate the counts $C(n,v,r)$

Evaluating Selectional Preference Models

- pseudo-word-evaluation: concatenate object of verb with the word closest in frequency: *drive..car* → *drive..car-house* and then let system decide which is the correct object

- human preferences for a test set of verb argument pairs (rate their degree of plausibility)

Bibliography

(1) Dan Jurafsky and James Martin: Speech and Language Processing (3rd ed. draft, version Jan 2022); Online: https://web.stanford.edu/~jurafsky/slp3/ (URL, Oct 2022); this slide-set is especially based on chapter 19

Recommendations for Studying

- minimal approach:

  work with the slides and understand their contents! Think beyond instead of merely memorizing the contents

- standard approach:

  minimal approach + read the corresponding pages in Jurafsky [1]

- interested students

  == standard approach