CSE 517
Natural Language Processing
Winter 2015

Frames

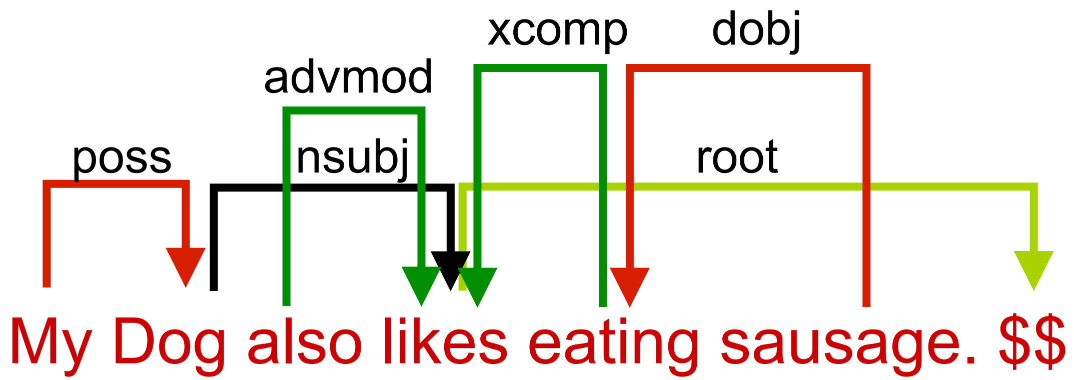
Yejin Choi

Some slides adapted from Martha Palmer, Chris Manning, Ray Mooney, Lluis Marquez ...

Overview

- Dependency Tree (very briefly)
- Selectional Preference
- Frames

Dependency structure



- Words are linked from head to dependent
- Warning! Some people do the arrows one way; some the other way
- Usually add a fake ROOT so every word is a dependent

- The idea of dependency structure goes back a long way
  - To Pāṇini's grammar (c. 5th century BCE)
- Constituency is a new-fangled invention
  - 20th century invention

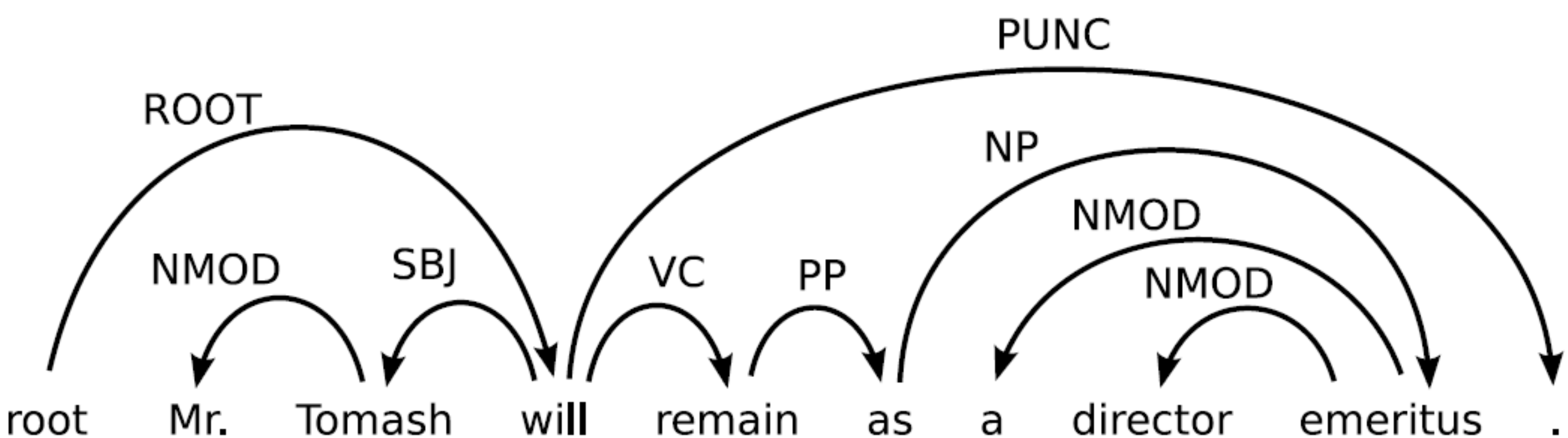
Relation between CFG to dependency parse

- Head
  - A dependency grammar has a notion of a head
  - Officially, CFGs don’t
  - But modern linguistic theory and all modern statistical parsers (Charniak, Collins, Stanford, …) do, via hand-written phrasal “head rules”:
- Conversion between CFG and Dependency Tree
  - The head rules can be used to extract a dependency parse from a CFG parse (follow the heads).
  - The extracted dependencies might not be correct (non-projective dependencies cannot be read off from CFG)
  - A phrase structure tree can be obtained from a dependency tree, but dependents are flat (no VP!)

Projective Dependencies

- Projective dependencies: when the tree edges are drawn directly on a sentence, it forms a tree (without a cycle), and there is no crossing edge.

- Projective Dependency:
- Eg:

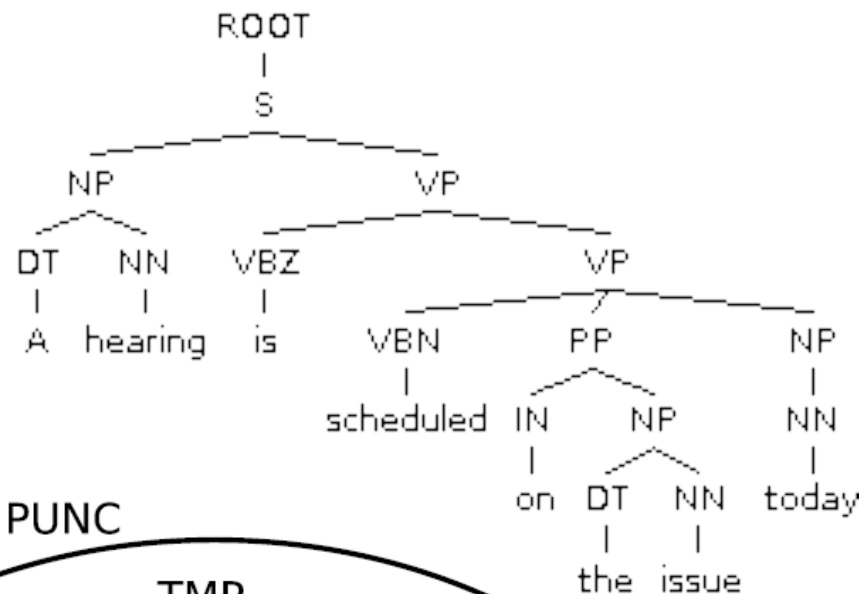


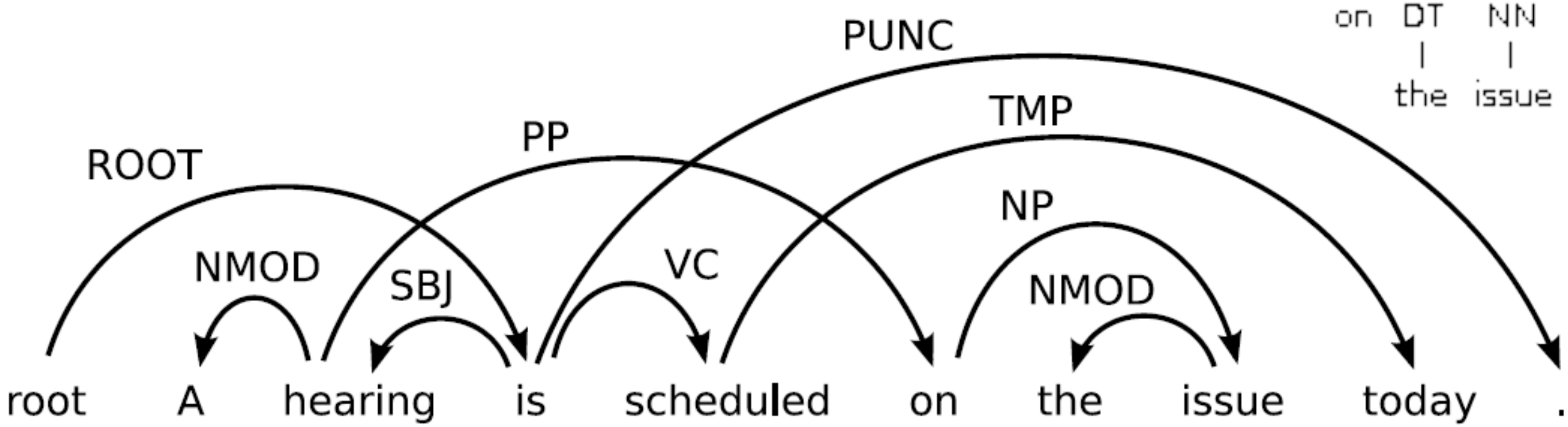
Example from Mcdonald and Satta (2007)

Non Projective Dependencies

- Non-Projective dependencies contain:
  - cycles

  - crossing edges



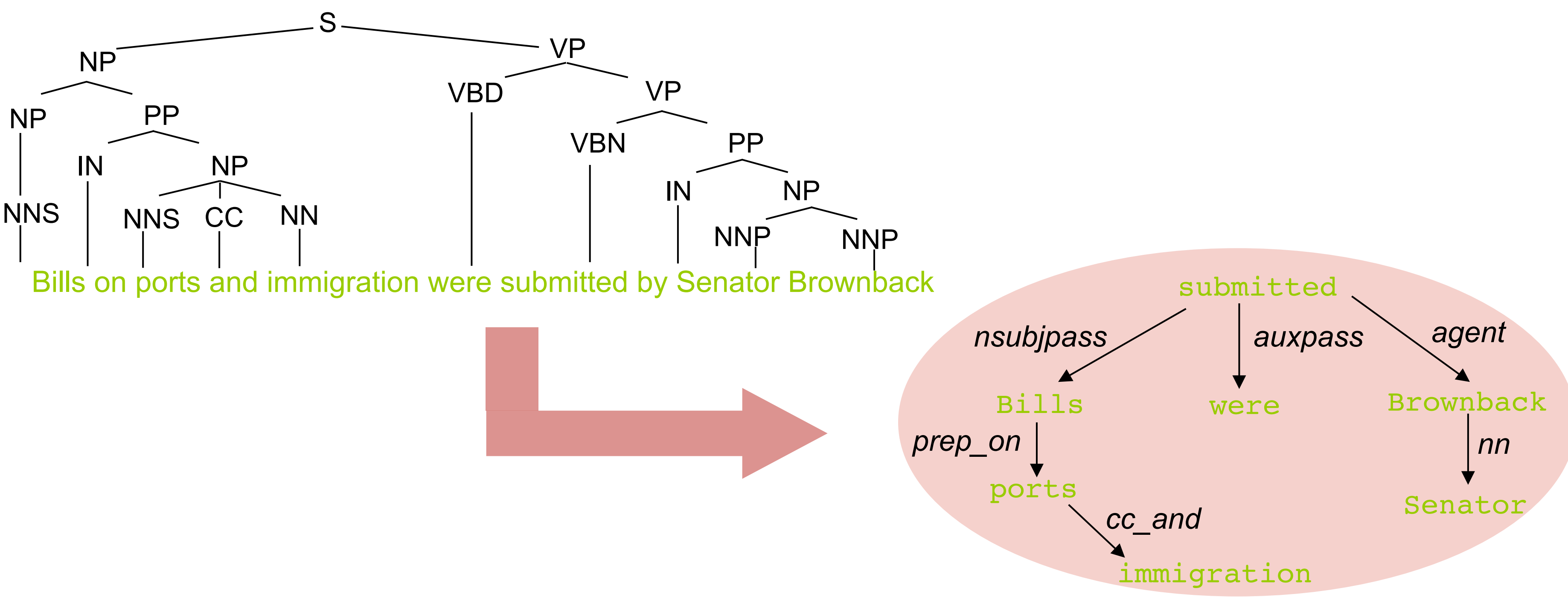


Example from Mcdonald and Satta (2007)

Extracting grammatical relations from statistical constituency parsers

[de Marneffe et al. LREC 2006]

- Exploit the high-quality syntactic analysis done by statistical constituency parsers to get the grammatical relations [typed dependencies]
- Dependencies are generated by pattern-matching rules



Grammatical Roles

- Dependency relations closely relate to grammatical roles
- Argument Dependencies
  - nsubj – nominal subject
  - nsubjpass – nominal subject in passive voice
  - dobj – direct object
  - pobj – object of preposition
- Modifier Dependencies
  - det – determiner
  - prep – prepositional modifier
  - mod
- Online Demos:
  - Stanford parser: http://nlp.stanford.edu:8080/parser/
  - Turbo parser: http://demo.ark.cs.cmu.edu/parse

Overview

- Dependency Tree
- Selectional Preference
- Frames

Selectional Preference

- Semantic relations between predicates -- arguments
- Selectional Restriction:
  - semantic type constraint a predicate imposes on its arguments --- certain semantic types are not allowed
  - I want to eat someplace that's close to school.
    - => "eat" is intransitive
  - I want to eat Malaysian food.
    - => "eat" is transitive
  - "eat" expects its object to be edible (when the subject is an animate).
- Selectional Preference:
  - Preferences among allowed semantic types
  - [a living entity] eating [food]
  - [concerns, zombies, ...] eating [a person]

Selectional Preference

- Some words have stronger selectional preference than others
  - imagine ...
  - diagonalize ...
- P(C) := the distribution of semantic classes (concepts)
- P(C|v) := the distribution of semantic classes of the object of the given verb 'v'
  - What does it mean if P(C) = P(C|v) ?
- How to quantify the distance between two distributions?
  - Kullback-Leibler divergence (KL divergence)

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

Selectional Preference

- Selectional preference of a predicate 'v':

$$S(v) = D(P(C|v)||P(C)) = \sum_c P(c|v) \log \frac{P(c|v)}{P(c)}$$

- Selectional association between 'v' and 'c' (Resnik 1996)

$$A(v,c) = \frac{1}{S(v)} P(c|v) \log \frac{P(c|v)}{P(c)}$$

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

- KL Divergence $D(P||G) = \sum_x P(x) \log \frac{P(x)}{Q(x)}$

Overview

- Dependency Tree
- Selectional Preference
- Frames

Frames

“Case for Case”

- Theory:
  - Frame Semantics (Fillmore 1968)
- Resources:
  - VerbNet(Kipper et al., 2000)
  - FrameNet (Fillmore et al., 2004)
  - PropBank (Palmer et al., 2005)
  - NomBank
- Statistical Models:
  - Task: Semantic Role Labeling (SRL)

Frame Semantics

- Frame: Semantic frames are schematic representations of situations involving various participants, props, and other conceptual roles, each of which is called a frame element (FE)
- These include events, states, relations and entities.

✓ **Frame**: *"The case for case"* (Fillmore 1968)

- 8k citations in Google Scholar!

✓ **Script:** knowledge about situations like eating in a restaurant.

- "*Scripts, Plans, Goals and Understanding: an Inquiry into Human Knowledge Structures*" (Schank & Abelson 1977)

✓ **Political Framings**: George Lakoff's recent writings on the framing of political discourse.

C4C: Capturing Generalizations over Related Predicates & Arguments

VERB	BUYER	GOODS	SELLER	MONEY	PLACE
buy	subject	object	from	for	at
sell	to				
cost	indirect object	subject		object	at
spend	subject	on		object	at

Example from Ken Church (at Fillmore tribute workshop)

Case Grammar -> Frames

- Valency: Predicates have arguments (optional & required)
  - Example: “give” requires 3 arguments:
    - Agent (A), Object (O), and Beneficiary (B)
    - Jones (A) gave money (O) to the school (B)
- Frames:
  - commercial transaction frame: Buy/Sell/Pay/Spend
  - Save <good thing> from <bad situation>
  - Risk <valued object> for <situation>|<purpose>|<beneficiary>|<motivation>
- Collocations & Typical predicate argument relations
  - Save whales from extinction (not vice versa)
  - Ready to risk everything for what he believes
- Representation Challenges: What matters for practical NLP?
  - POS? Word order? Frames (typical predicate – arg relations)?

Slide from Ken Church (at Fillmore tribute workshop)

Thematic (Semantic) Roles

- AGENT - the volitional causer of an event
  - The waiter spilled the soup
- EXPERIENCER - the experiencer of an event
  - John has a headache
- FORCE - the non-volitional causer of an event
  - The wind blows debris from the mall into our yards.
- THEME - the participant most directly affected by an event
  - Only after Benjamin Franklin broke the ice ...
- RESULT - the end product of an event
  - The French government has built a regulation-size baseball diamond ...

Thematic (Semantic) Roles

- INSTRUMENT - an instrument used in an event
  - He turned to poaching catfish, stunning them with a shocking device ...
- BENEFICIARY - the beneficiary of an event
  - Whenever Ann makes hotel reservations for her boss ...
- SOURCE - the origin of the object of a transfer event
  - I flew in from Boston
- GOAL - the destination of an object of a transfer event
  - I drove to Portland

- Can we read semantic roles off from PCFG or dependency parse trees?

Semantic roles ≠ Grammatical roles

- Agent – the volitional causer of an event
  - usually “subject”, sometimes “prepositional argument”, ...
- Theme – the participant directly affected by an event
  - usually “object”, sometimes “subject”, ...
- Instrument – an instrument (method) used in an event
  - usually prepositional phrase, but can also be a “subject”

- John broke the window.
- John broke the window with a rock.
- The rock broke the window.
- The window broke.
- The window was broken by John.

Ergative Verbs

- **Ergative verbs**
  - **subject when intransitive** = **direct object when transitive**.
  - "it broke the window" (transitive)
  - "the window broke" (intransitive).
- Most verbs in English are *not* ergative (the subject role does not change whether transitive or not)
  - "He ate the soup" (transitive)
  - "He ate" (intransitive)
- Ergative verbs generally describe some sort of “changes” of states:
  - Verbs suggesting a change of state — *break, burst, form, heal, melt, tear, transform*
  - Verbs of cooking — *bake, boil, cook, fry*
  - Verbs of movement — *move, shake, sweep, turn, walk*
  - Verbs involving vehicles — *drive, fly, reverse, run, sail*

FrameNet

Words in "change_position_on _a_scale" frame:

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	

- Frame := the set of words sharing a similar predicate-argument relations
- Predicate can be a verb, noun, adjective, adverb
- The same word with multiple senses can belong to multiple frames

Roles in "**change_position_on _a_scale**" frame

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.

Example

ATTRIBUTE
DIFFERENCE
FINAL_STATE
FINAL_VALUE
INITIAL_STATE
INITIAL_VALUE
ITEM
VALUE_RANGE

DURATION
SPEED
GROUP

- [Oil] rose [in price] [by 2%].
- [It] has increased [to having them 1 day a month].
- [Microsoft shares] fell [to 7 5/8].
- [cancer incidence] fell [by 50%] [among men].
- a steady increase [from 9.5] [to 14.3] [in dividends].
- a [5%] [dividend] increase…

Find “Item” roles?

ATTRIBUTE
DIFFERENCE
FINAL_STATE
FINAL_VALUE
INITIAL_STATE
INITIAL_VALUE
ITEM
VALUE_RANGE
DURATION
SPEED
GROUP

- [Oil] rose [in price] [by 2%].
- [It] has increased [to having them] [1 day a month].
- [Microsoft shares] fell [to 7 5/8].
- [cancer incidence] fell [by 50%] [among men].
- a steady increase [from 9.5] [to 14.3] [in dividends].
- a [5%] [dividend] increase…

Find “Difference” & “Final_Value” roles?

ATTRIBUTE DIFFERENCE FINAL_STATE FINAL_VALUE INITIAL_STATE INITIAL_VALUE ITEM VALUE_RANGE
DURATION SPEED GROUP

- [Oil] rose [in price] [by 2%].
- [It] has increased [to having them] [1 day a month].
- [Microsoft shares] fell [to 7 5/8].
- [cancer incidence] fell [by 50%] [among men].
- a steady increase [from 9.5] [to 14.3] [in dividends].
- a [5%] [dividend] increase…

FrameNet (2004)

- Project at UC Berkeley led by Chuck Fillmore for developing a database of frames, general semantic concepts with an associated set of roles.
- Roles are specific to frames, which are “invoked” by the predicate, which can be a verb, noun, adjective, adverb
  - JUDGEMENT frame
    - Invoked by: V: blame, praise, admire; N: fault, admiration
    - Roles: JUDGE, EVALUEE, and REASON
- Specific frames chosen, and then sentences that employed these frames selected from the British National Corpus and annotated by linguists for semantic roles.
- Initial version: 67 frames, 1,462 target words,
  49,013 sentences, 99,232 role fillers

PropBank
(proposition bank)

PropBank := proposition bank (2005)

- Project at Colorado lead by Martha Palmer to add semantic roles to the Penn treebank.
- Proposition := verb + a set of roles
- Annotated over 1M words of Wall Street Journal text with existing gold-standard parse trees.
- Statistics:
  - 43,594 sentences 99,265 propositions
  - 3,324 unique verbs 262,281 role assignments

PropBank argument numbering

- Numbered roles, rather than named roles.
  - Arg0, Arg1, Arg2, Arg3, …
- Different numbering scheme for each **verb sense**.
- The general pattern of numbering is as follows.

- **Arg0** = "Proto-Agent" (agent)
- **Arg1** = "Proto-Patient" (direct object / theme / patient)
- Arg2 = indirect object (benefactive / instrument / attribute / end state)
- Arg3 = start point (benefactive / instrument / attribute)
- Arg4 = end point

Different “frameset” for each verb sense

- Mary left the room.
- Mary left her daughter-in-law her pearls in her will.

Frameset **leave.01** "move away from":
Arg0: entity leaving
Arg1: place left

Frameset **leave.02** "give":
Arg0: giver
Arg1: thing given
Arg2: beneficiary

PropBank argument numbering

Argument numbering conserving the common semantic roles shared among predicates that belong to a related frame

Buy	Sell
Arg0: buyer	Arg0: seller
Arg1: goods	Arg1: goods
Arg2: seller	Arg2: buyer
Arg3: rate	Arg3: rate
Arg4: payment	Arg4: payment

Ergative Verbs

Sales <u>rose</u> 4% to $3.28 billion from $3.16 billion.

The Nasdaq composite index <u>added</u> 1.01 to 456.6 on paltry volume.

Semantic Roles (per PropBank)

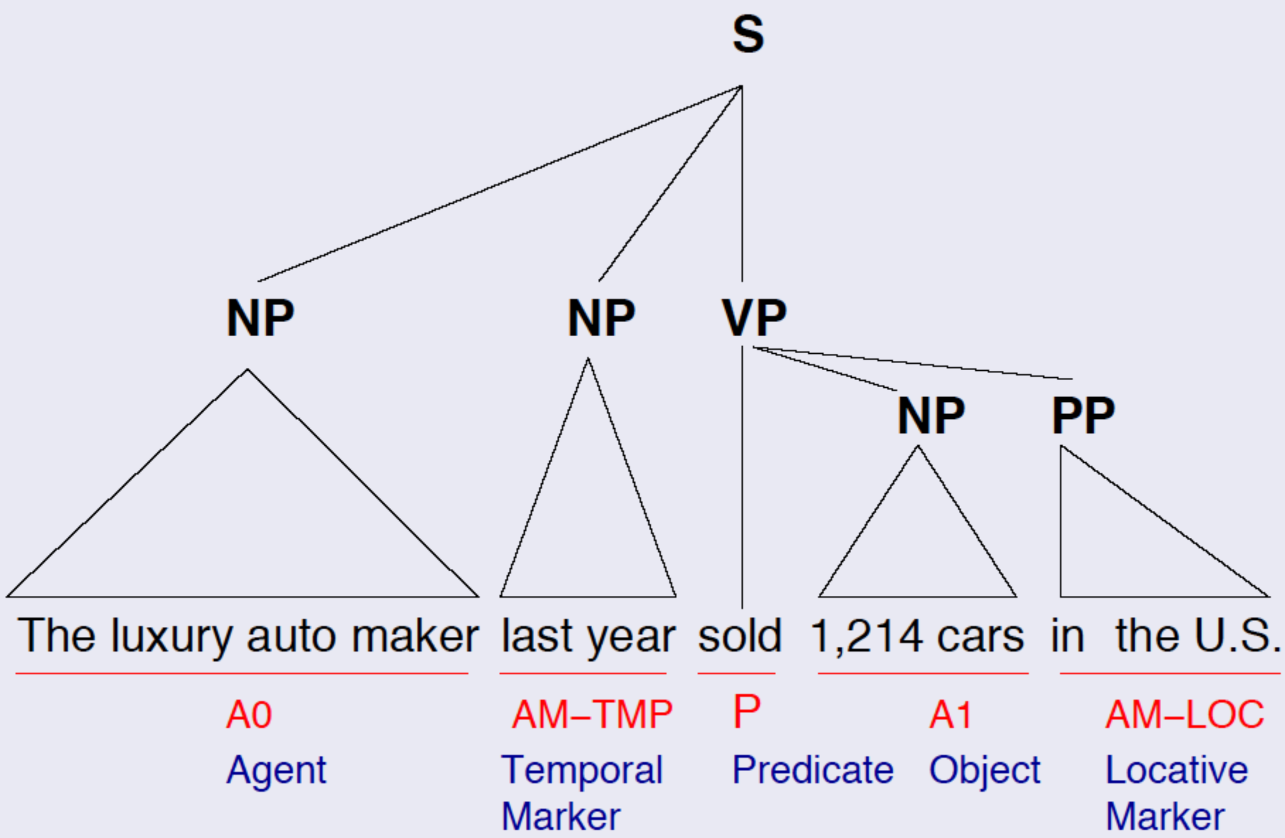
- Arg0 = None (unaccusative, i.e, no agent)
- Arg1 = patient, thing rising
- Arg2 = amount risen
- Arg3 = start point
- Arg4 = end point

Semantic Role Labeling

Semantic Role Labeling (Task)

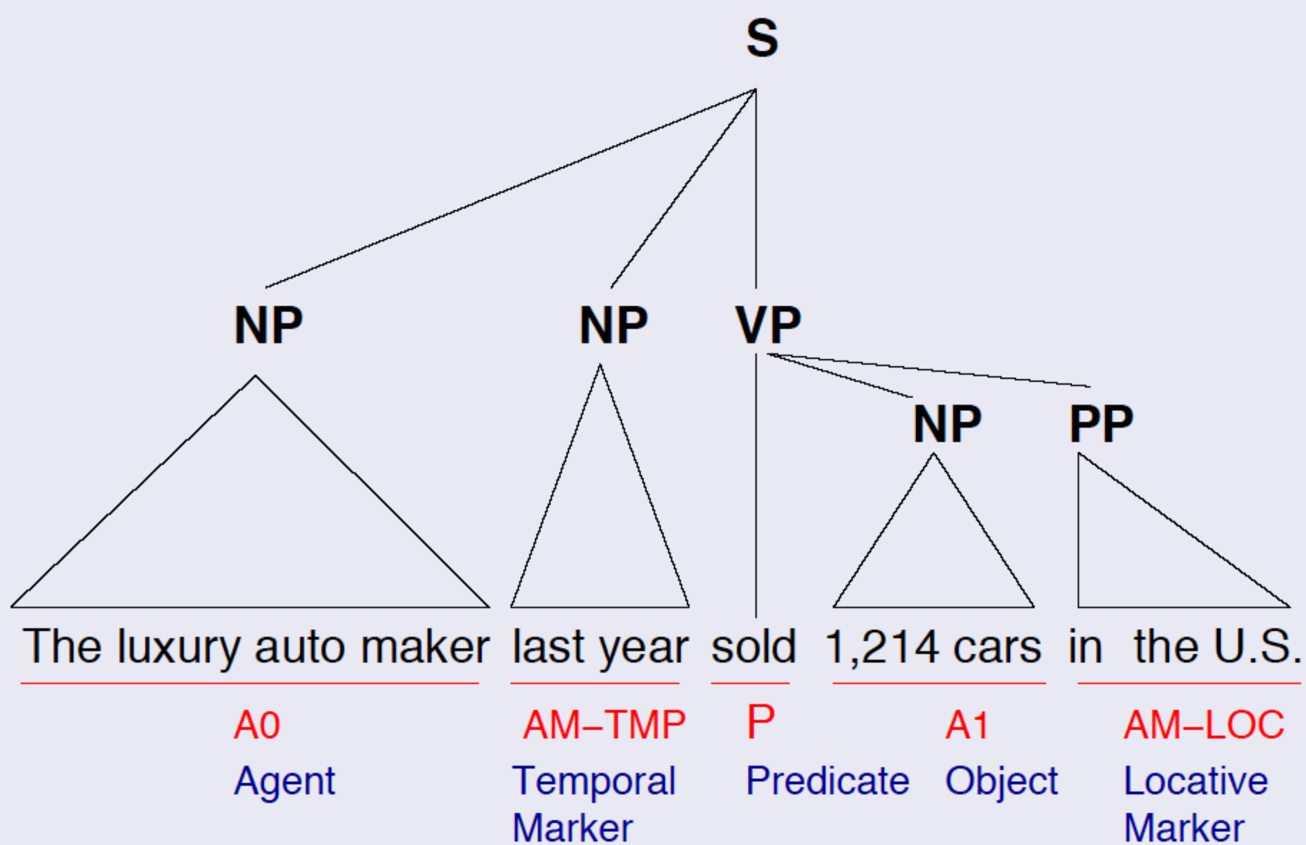
- Shallow meaning representation beyond syntactic parse trees
- Question Answering
  - “Who” questions usually use Agents
  - “What” question usually use Patients
  - “How” and “with what” questions usually use Instruments
  - “Where” questions frequently use Sources and Destinations.
  - “For whom” questions usually use Beneficiaries
  - “To whom” questions usually use Destinations
- Machine Translation Generation
  - Semantic roles are usually expressed using particular, distinct syntactic constructions in different languages.
- Summarization, Information Extraction

SRL $\stackrel{def}{=}$ detecting basic event structures such as *who* did *what* to *whom*, *when* and *where* [**IE** point of view]



Example from Lluis Marquez

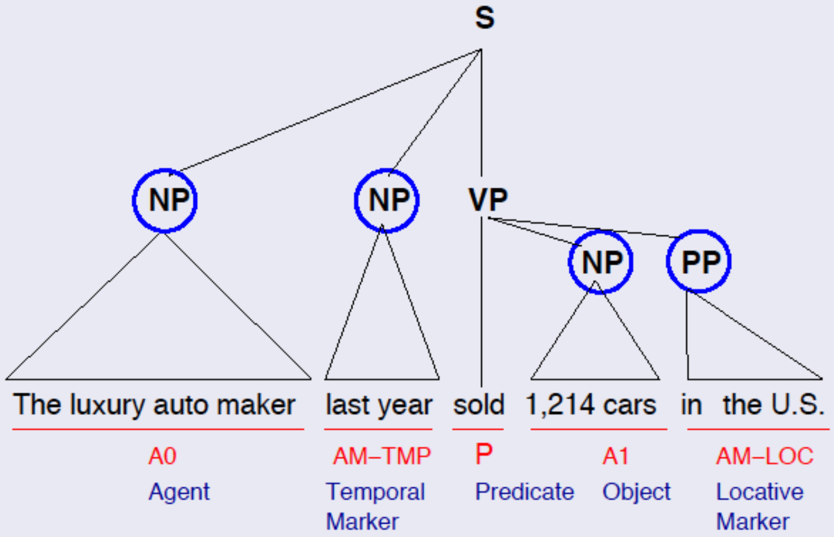
SRL $\stackrel{def}{=}$ identify the *arguments* of a given verb and assign them semantic labels describing the *roles* they play in the predicate (i.e., identify predicate argument structures) [**CL** point of view]



Example from Lluis Marquez

Linguistic nature of the problem

- Argument identification is strongly related to syntax



- Role labeling is a semantic task
  - e.g., selectional preferences should play an important role

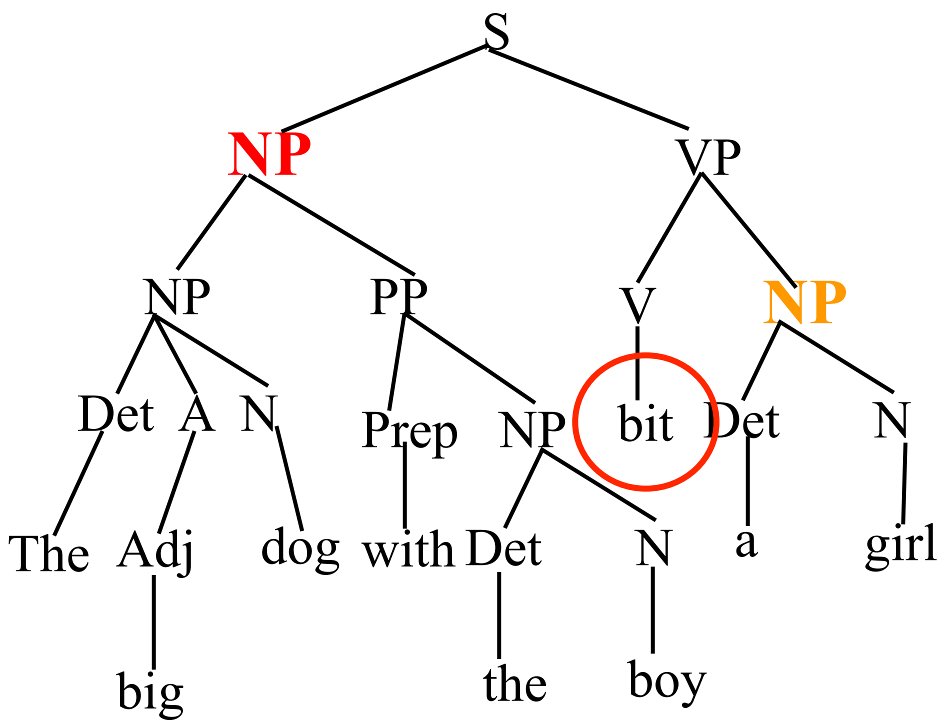
Example from Lluis Marquez

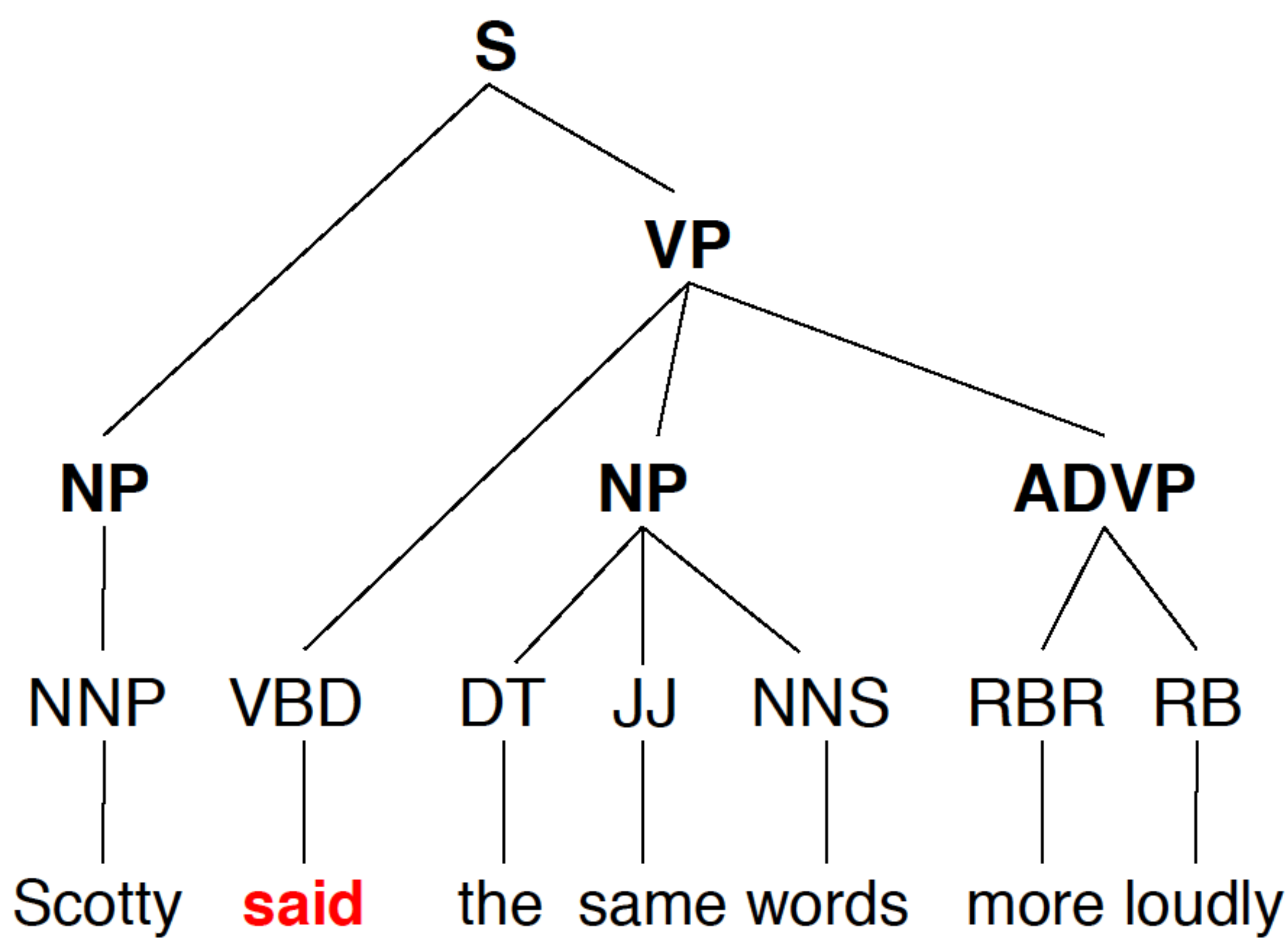
SRL as Parse Node Classification

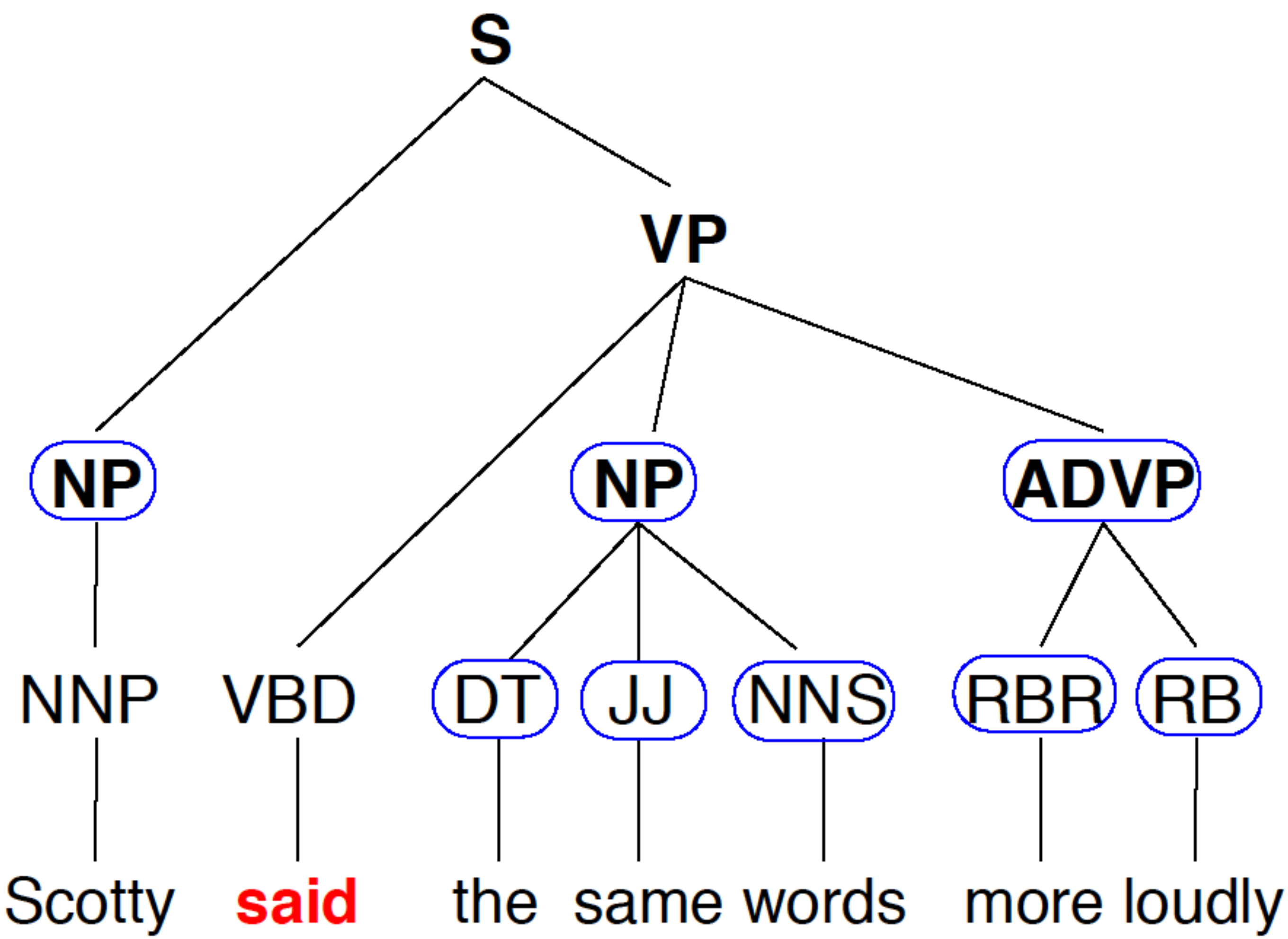
- Assume that a syntactic parse is available
- Treat problem as classifying parse-tree nodes.
- Can use any machine-learning classification method.
- Critical issue is engineering the right set of features for the classifier to use.

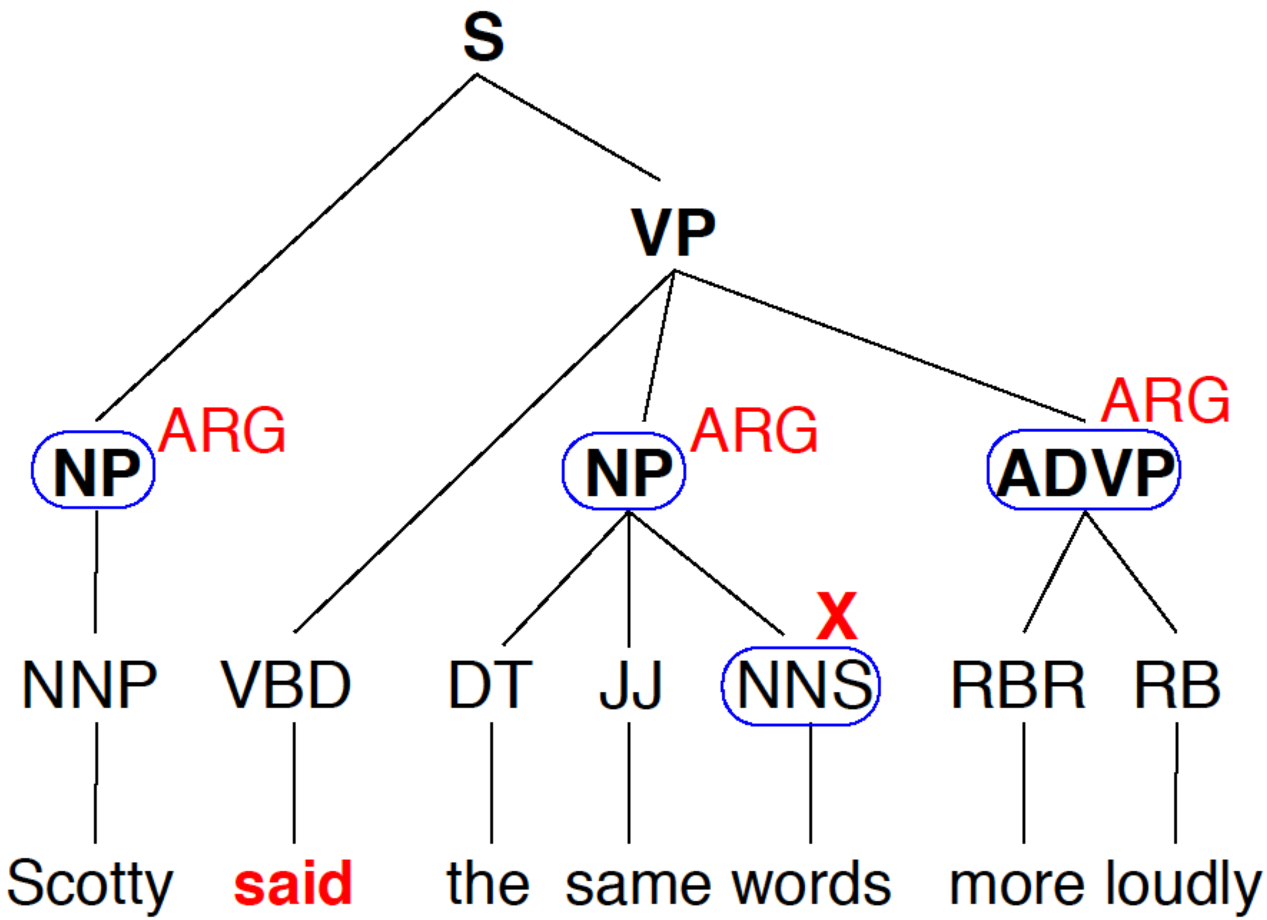
Color Code:

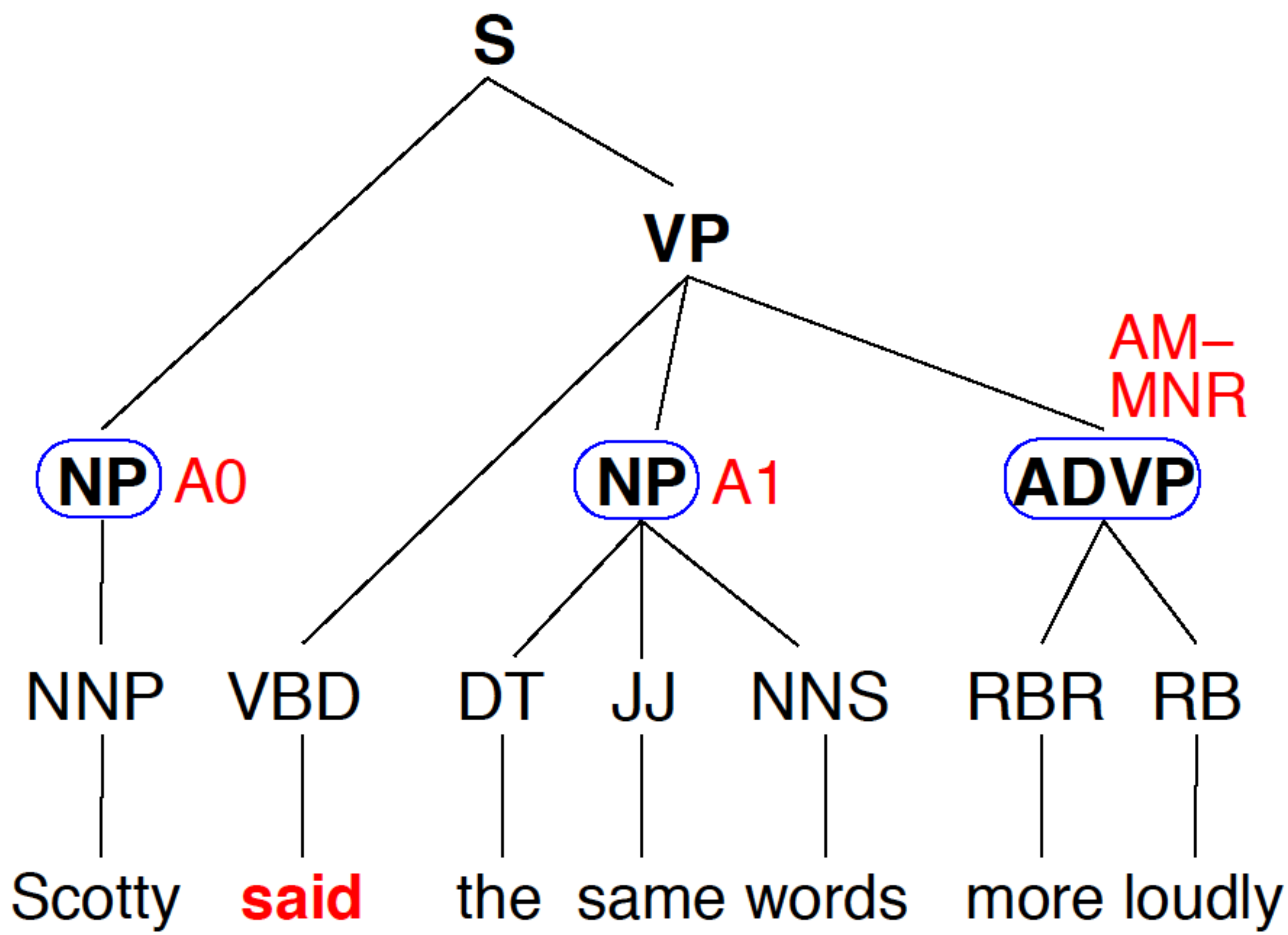
not-a-role
agent
patient
source
destination
instrument
beneficiary

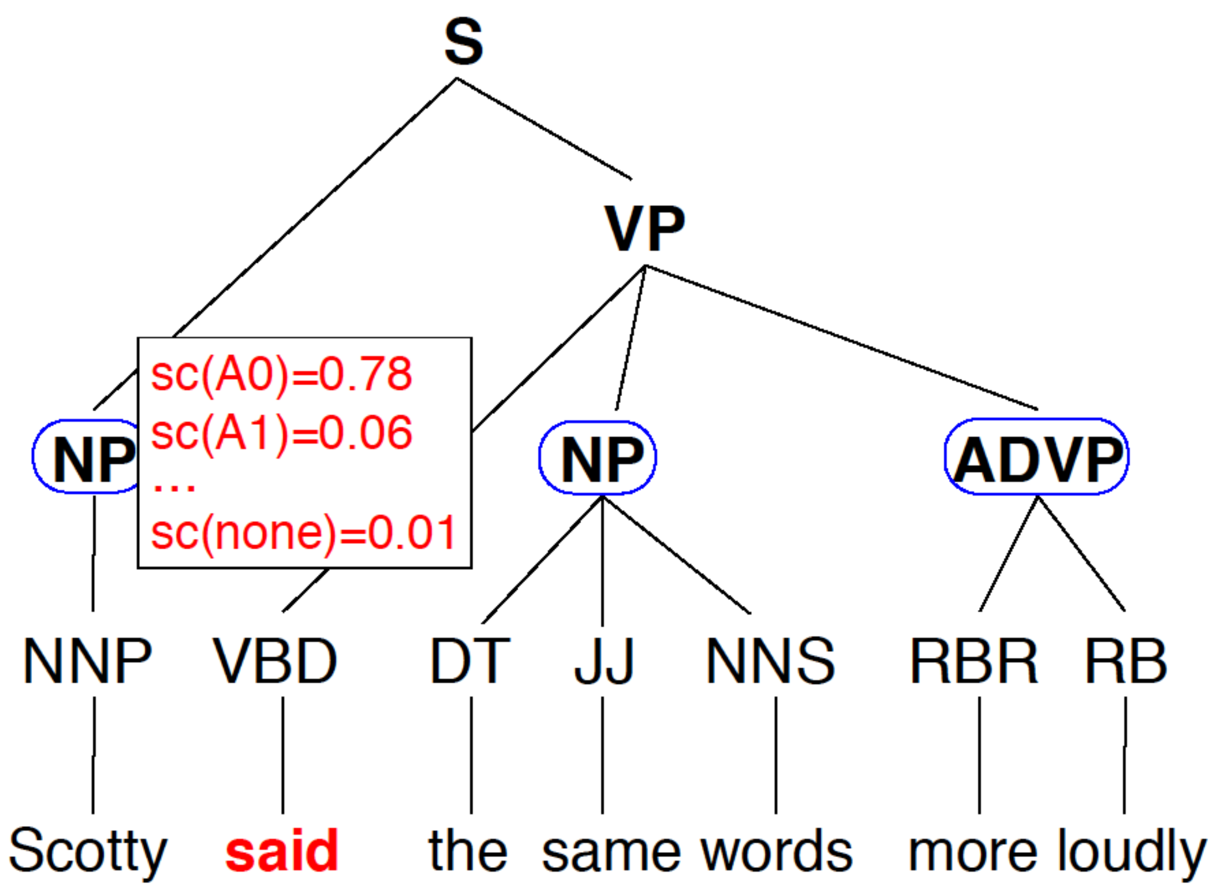


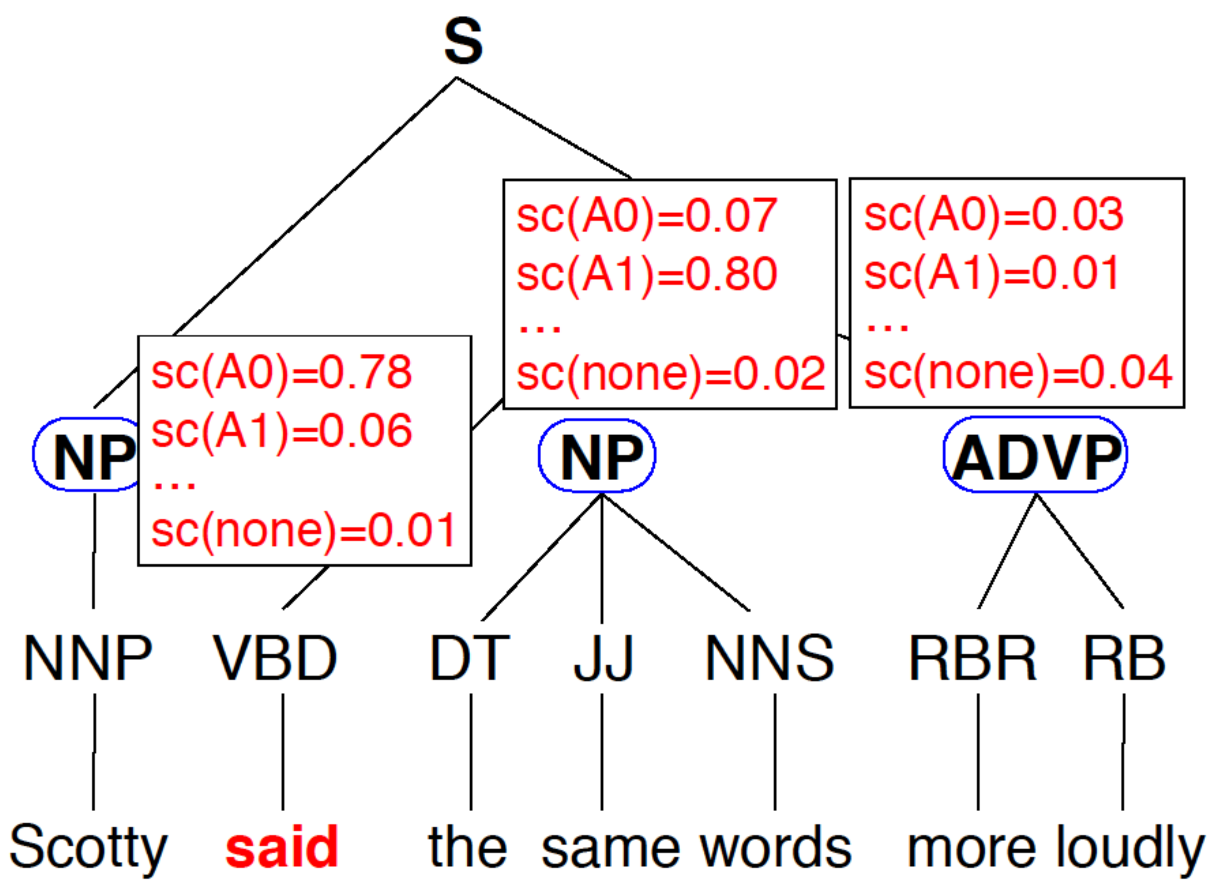

S
VP
NP
NP
ADVP
NNP
VBD
DT
JJ
NNS
RBR
RB
Scotty
said
the
same
words
more
loudly

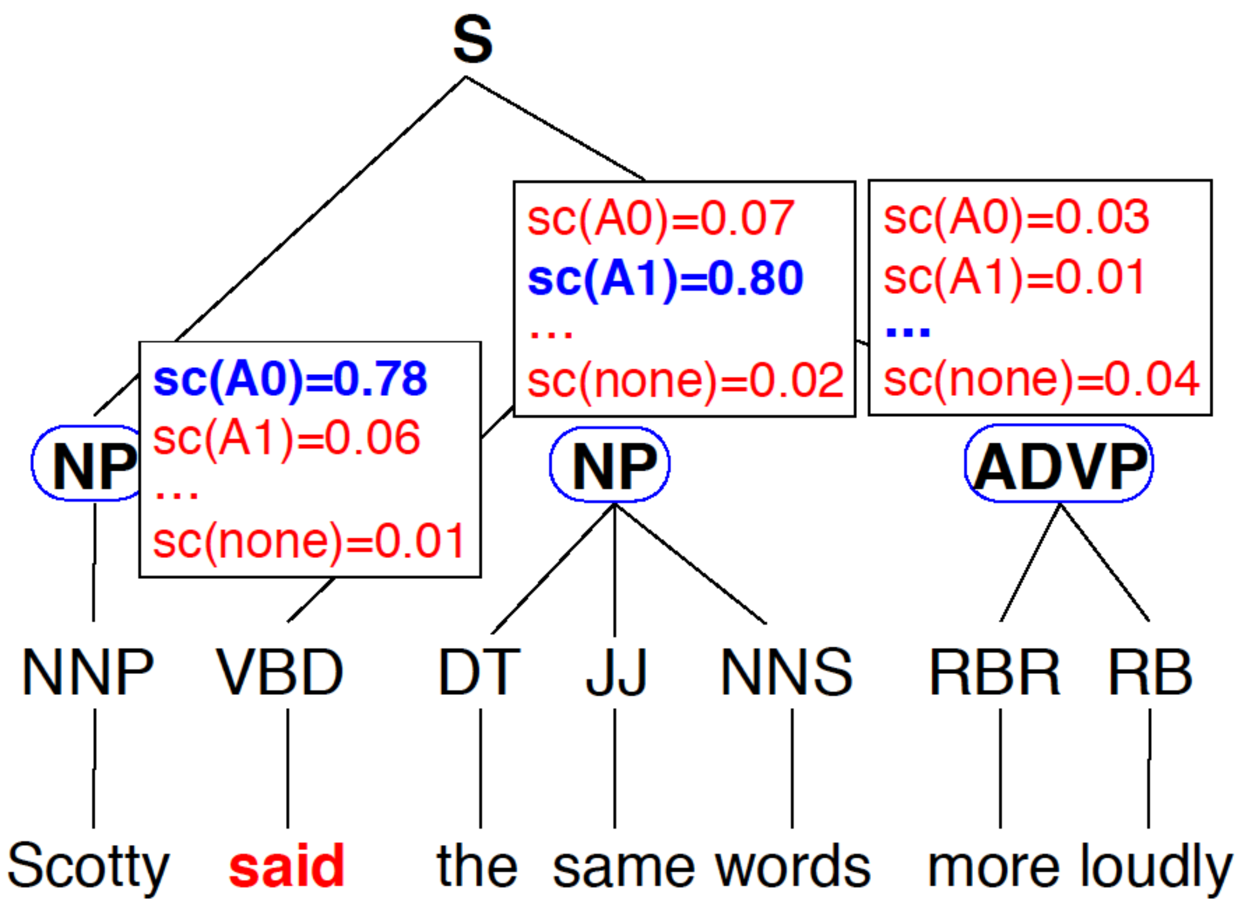

S
VP
NP
NP
ADVP
NNP
VBD
DT
JJ
NNS
RBR
RB
Scotty
said
the
same
words
more
loudly


S
VP
NP
ARG
NP
ARG
ADVP
ARG
X
NNP
VBD
DT
JJ
NNS
RBR
RB
Scotty
said
the
same
words
more
loudly


S
VP
AM–
MNR
NP A0
NP A1
ADVP
NNP VBD DT JJ NNS RBR RB
Scotty said the same words more loudly

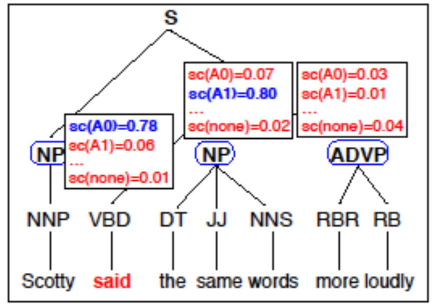

S
VP
NP
sc(A0)=0.78
sc(A1)=0.06
...
sc(none)=0.01
NP
ADVP
NNP
VBD
DT
JJ
NNS
RBR
RB
Scotty
said
the
same
words
more
loudly


S
sc(A0)=0.07
sc(A1)=0.80
...
sc(none)=0.02
sc(A0)=0.03
sc(A1)=0.01
...
sc(none)=0.04
sc(A0)=0.78
sc(A1)=0.06
...
sc(none)=0.01
NP
NP
ADVP
NNP
VBD
DT
JJ
NNS
RBR
RB
Scotty
said
the
same
words
more
loudly

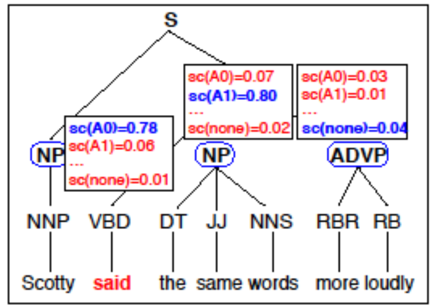

S
sc(A0)=0.07
sc(A1)=0.80
...
sc(none)=0.02
sc(A0)=0.03
sc(A1)=0.01
...
sc(none)=0.04
sc(A0)=0.78
sc(A1)=0.06
...
sc(none)=0.01
NP
NP
ADVP
NNP
VBD
DT
JJ
NNS
RBR
RB
Scotty
said
the
same
words
more
loudly

Issues in Parse Node Classification

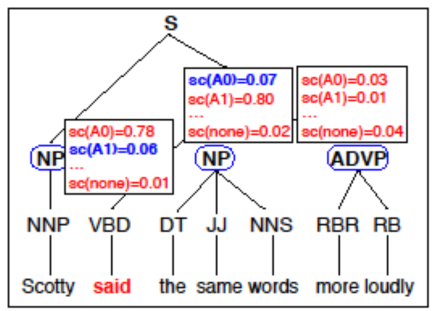
- Results may violate constraints like “an action has at most one agent”?
  - Use some method to enforce constraints when making final decisions. i.e. determine the most likely assignment of roles that also satisfies a set of known constraints.
- Due to errors in syntactic parsing, the parse tree is likely to be incorrect.
  - Try multiple top-ranked parse trees and somehow combine results.
  - Integrate syntactic parsing and SRL.



GS = 0.45

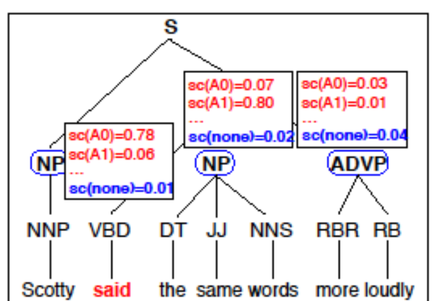


GS = 0.15



GS = 0.09

...



GS = 0.001

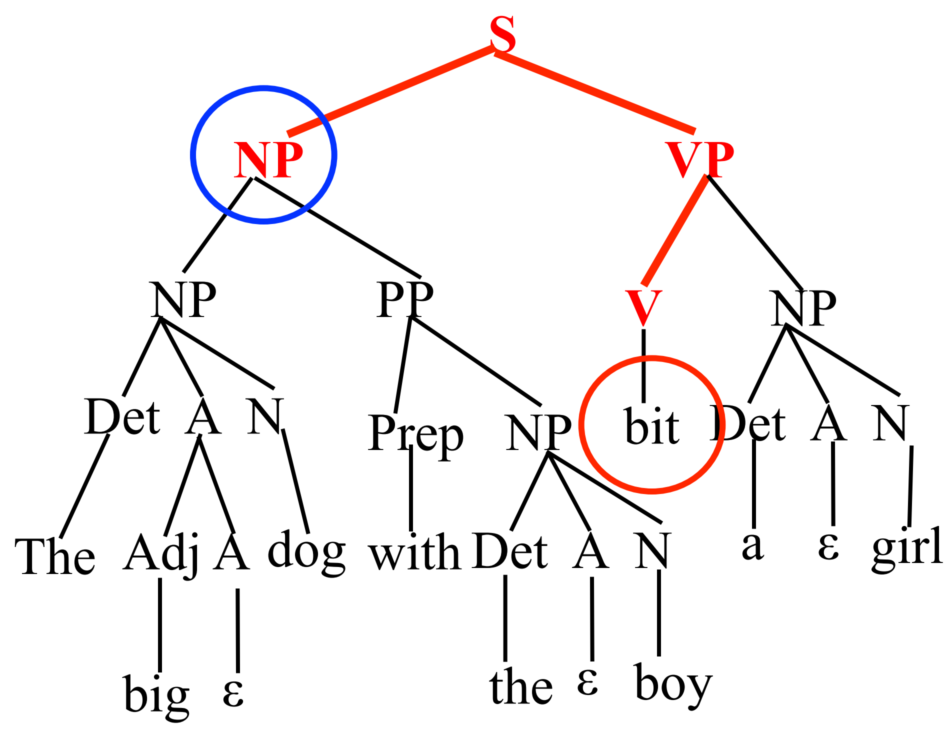
Syntactic Features for SRL

- **Phrase type**: The syntactic label of the candidate role filler (e.g. NP).
- **Parse tree path**: The path in the parse tree between the predicate and the candidate role filler.

Parse Tree Path Feature: Example 1

Path Feature Value:

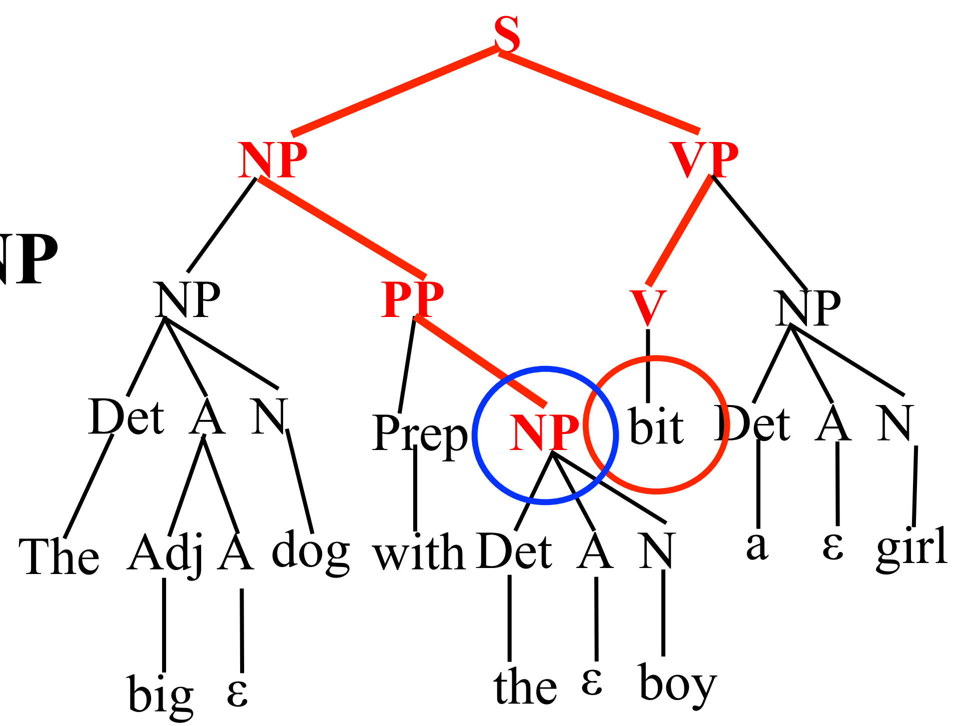
V ↑ VP ↑ S ↓ NP



Parse Tree Path Feature: Example 2

Path Feature Value:

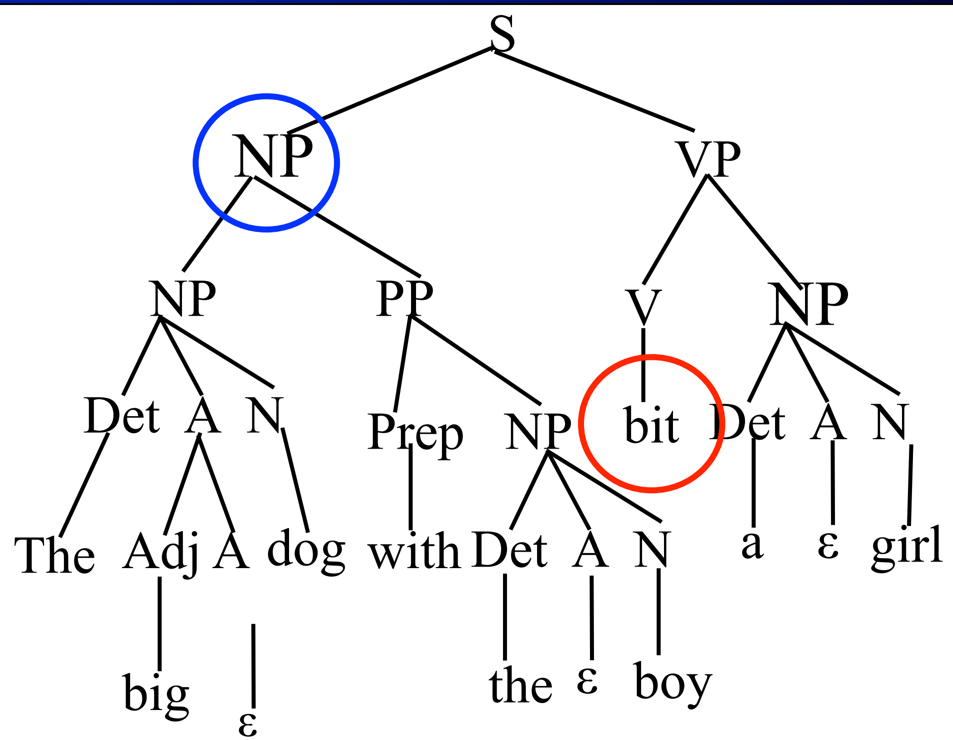
V ↑ VP ↑ S ↓ NP ↓ PP ↓ NP



Features for SRL

- **Phrase type**: The syntactic label of the candidate role filler (e.g. NP).
- **Parse tree path**: The path in the parse tree between the predicate and the candidate role filler.
- **Position**: Does candidate role filler ***precede*** or ***follow*** the predicate in the sentence?
- **Voice**: Is the predicate an ***active*** or ***passive*** verb?
- **Head Word**: What is the head word of the candidate role filler?

Features for SRL



Phrase type	Parse Path	Position	Voice	Head word
NP	**V↑VP↑S↓NP**	precede	active	dog

Selectional Preference

- Selectional preference/restrictions are constraints that certain verbs place on the filler of certain semantic roles.
  - Agents should be animate
  - Beneficiaries should be animate
  - Instruments should be tools
  - Patients of “eat” should be edible
  - Sources and Destinations of “go” should be places.
  - Sources and Destinations of “give” should be animate.
- Taxanomic abstraction hierarchies or ontologies (e.g. hypernym links in WordNet) can be used to determine if such constraints are met.
  - “John” is a “Human” which is a “Mammal” which is a “Vertebrate” which is an “Animate”

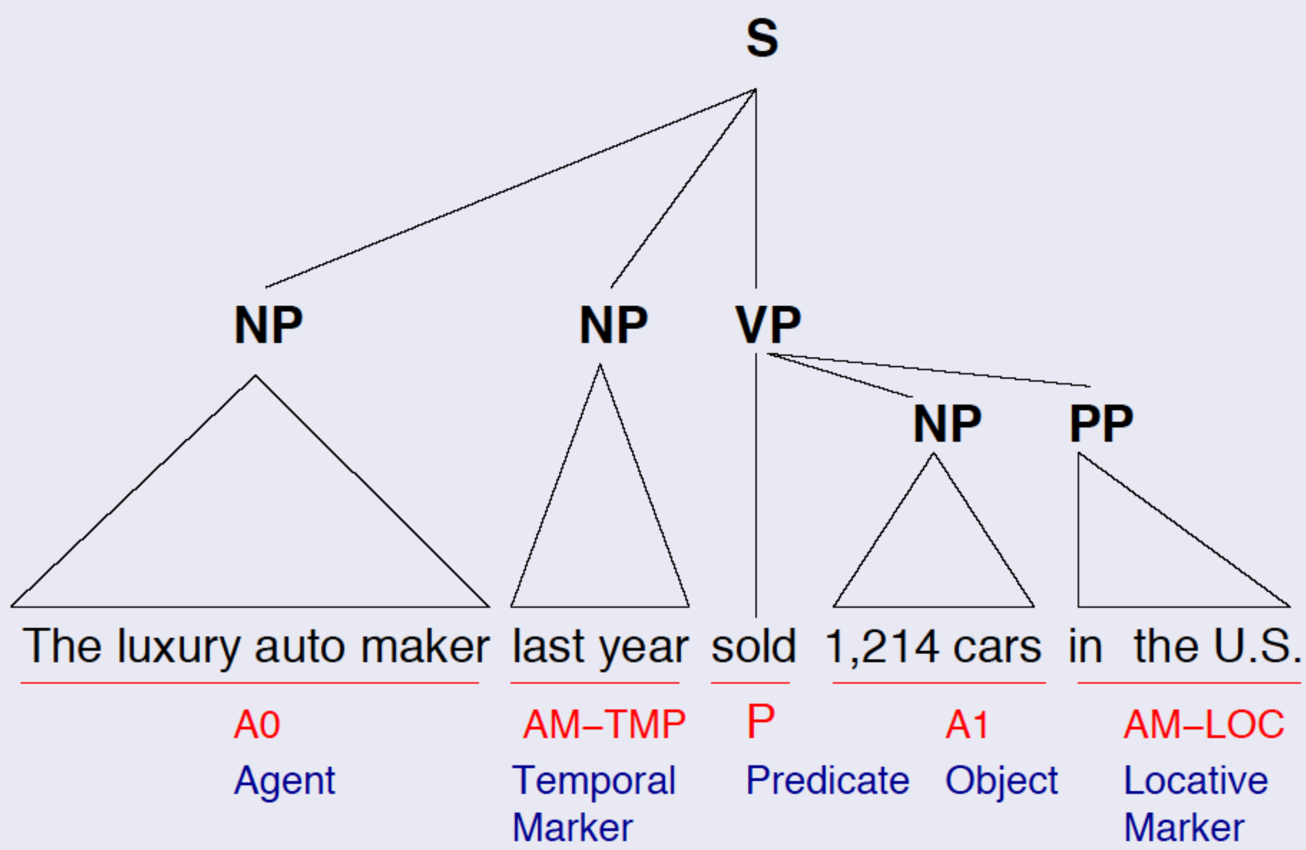
Selectional Preference & Syntactic Ambiguity

- Many syntactic ambiguities like PP attachment can be resolved using selectional restrictions.
  - "John ate the spaghetti with meatballs."
    "John ate the spaghetti with chopsticks."
    - Instruments should be tools
    - Patients of "eat" must be edible
  - "John hit the man with a dog."
    "John hit the man with a hammer."
    - Instruments should be tool

Use of Sectional Restrictions

- Selectional restrictions can help rule in or out certain semantic role assignments.
  - "John bought the car for $21K"
    - Beneficiaries should be Animate
    - Instrument of a "buy" should be Money
  - "John went to the movie with Mary"
    - Instrument should be Inanimate
  - "John drove Mary to school in the van"
    "John drove the van to work with Mary."
    - Instrument of a "drive" should be a Vehicle

SRL $\stackrel{def}{=}$ identify the *arguments* of a given verb and assign them semantic labels describing the *roles* they play in the predicate (i.e., identify predicate argument structures) [**CL** point of view]



Example from Lluis Marquez

When can we expect to learn frames?

- Corpus-size requirements:
  - freq(content words) ≈ parts per million
- 1970s Corpora: 1 M words (Brown Corpus)
  - Large enough to make a list of common content words
- 1990s: 100 M words (British National Corpus)
  - Large enough to see associations of common predicates with function words
    - "save" + "from"
  - Useful for parsing phrasal verbs: **V NP P** (Hindle & Rooth, 1993)
    - Most parsers are trained on Brown Corpus
    - (too small for phrasal verbs, let alone conjunction)
- Coming soon: $1M^2$ words (Google?)
  - Large enough to see associations of pairs of content words (collocations)
    - "give" + $$
    - "save" + "whale"
    - "save" + "extinction"
    - "risk" <valued object> for <purpose>
  - Useful for parsing every-way ambiguous Catalan Constructions (Church, 1980)
    - Conjunction, NN modification, PP attachment

Slide from Ken Church (at Fillmore tribute workshop)