CS460/626 : Natural Language Processing/Speech, NLP and the Web
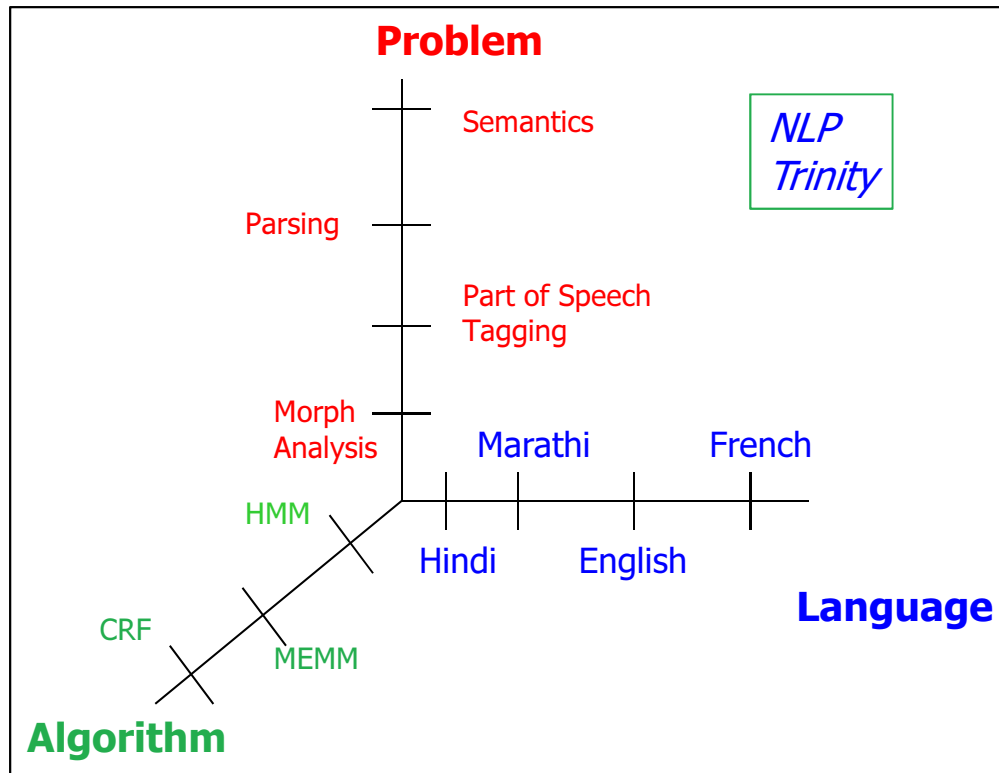
Lecture 23:
Binding Theory

Pushpak Bhattacharyya
CSE Dept.,
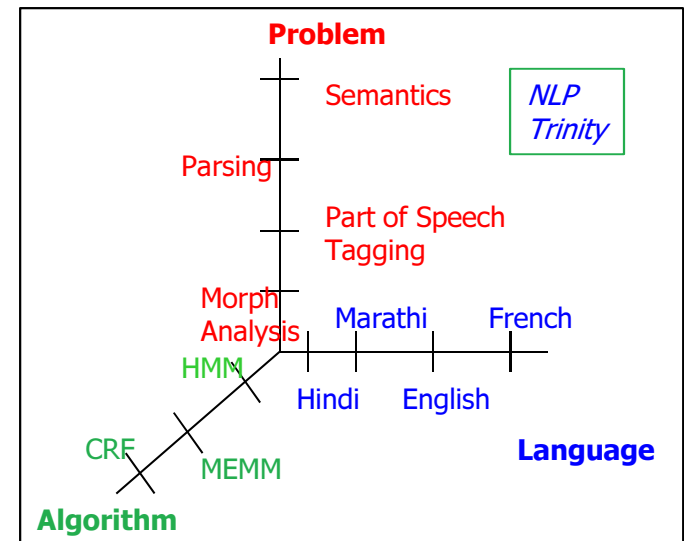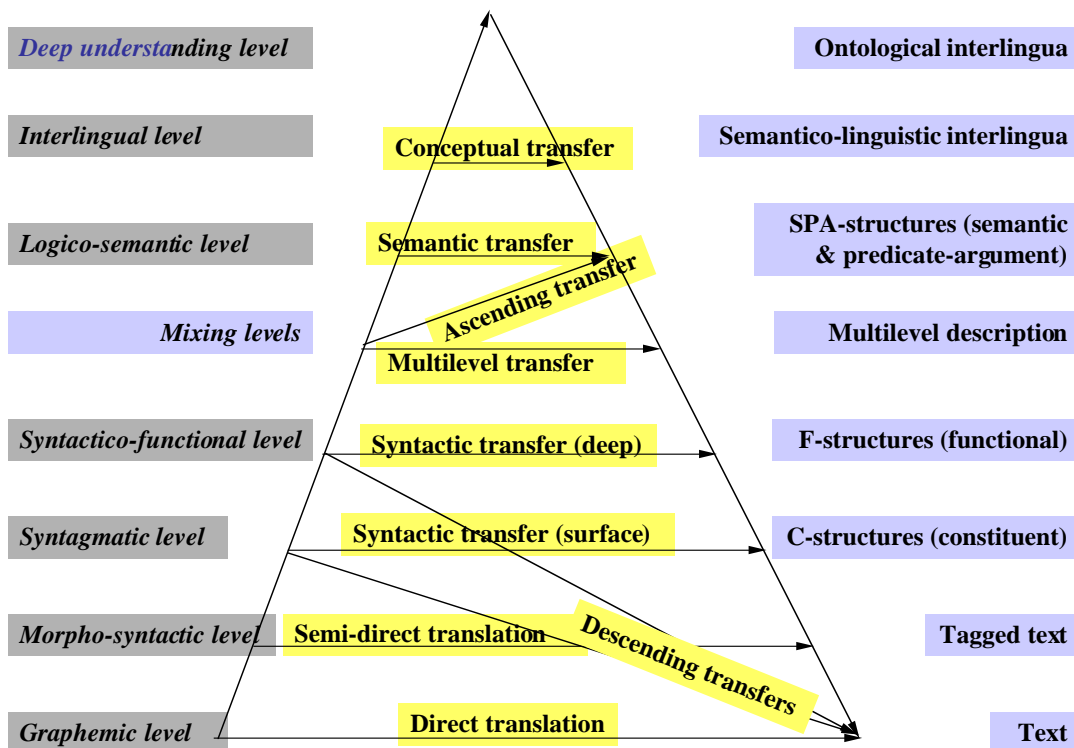IIT Bombay

8th Oct, 2012

Problem

Semantics

NLP Trinity

Parsing

Part of Speech Tagging

Morph Analysis

HMM

Marathi

French

CRF

Hindi

English

MEMM

Language

Algorithm

Coindexing

- *Ram$_p$ saw himself$_p$ in the mirror
- *Ram$_p$ saw himself$_q$ in the mirror
- Ram$_p$ saw him$_q$ in the mirror
- *Ram$_p$ saw him$_p$ in the mirror
- The grandmother$_p$ of Ram$_q$'s distant uncle$_r$ saw him$_s$ the mirror
- *The grandmother$_p$ of Ram$_q$'s distant uncle$_r$ saw him$_p$ the mirror
- The grandmother$_p$ of Ram$_q$'s distant uncle$_r$ saw herself$_p$ the mirror
- *The grandmother$_p$ of Ram$_q$'s distant uncle$_r$ saw himself$_p$ the mirror

Perspective

Level	Transfer	Structure
Deep understanding level		Ontological interlingua
Interlingual level	Conceptual transfer	Semantico-linguistic interlingua
Logico-semantic level	Semantic transfer	SPA-structures (semantic & predicate-argument)
Mixing levels	Multilevel transfer	Multilevel description
Syntactico-functional level	Syntactic transfer (deep)	F-structures (functional)
Syntagmatic level	Syntactic transfer (surface)	C-structures (constituent)
Morpho-syntactic level	Semi-direct translation	Tagged text
Graphemic level	Direct translation	Text

Ascending transfer

Descending transfers

Problem

Semantics

Parsing

Part of Speech Tagging

Morph Analysis

NLP Trinity

HMM

Marathi   French

Hindi   English

CRF

MEMM

Language

Algorithm

4

R-Expressions

- Consider the sentence: *Ram found a blanket in the green bag*
  - *Ram, a key in the green bag* are called "Referring Expressions" or "R-expressions"

- Key Definition: An **R-expression** is an entity that gets its meaning by referring to an entity in the world

Not all NPs are R-Expressions

- *Ram found himself a blanket in the green bag*
  - *Himself* must refer back to Ram and not to something in the outside world.

- Key Definition: An **Anaphor** is an NP that obligatorily gets its meaning from another NP in the sentence.

Types of Anaphors

- Reflexive pronouns
  - *Himself, herself, themselves*
- Reciprocals
  - Each other, one another
  - *Ram and Shyam saw each other*

Pronouns

- Get the meaning not necessarily from the same sentence

  - *Ram told Shyam that he should collect the blanket*

Pronouns with forward reference

- *That he will succeed, was known a priori to Ram*

Anaphors have definite syntactic position

- *himself Ram found a blanket*

Coindex and Antecedent

- Definition: An NP that gives its meaning to an anaphor (or pronoun) is called an **Antecedent**
- *Ram found himself a blanket*
  - *Ram:* antecedent
  - *Himself:* anaphor
- Coindexing convention:
  - *[Ram]$_i$ found [himself]$_i$ a [blanket]$_j$*
  - *[Ram]$_i$ told [Shyam]$_j$ that [he]$_i$ should collect the [blanket]$_k$*
  - *[Ram]$_i$ told [Shyam]$_j$ that [he]$_j$ should collect the [blanket]$_k$*
  - *[Ram]$_i$ told [Shyam]$_j$ that [he]$_k$ should collect the [blanket]$_l$*
- Definition: NPs with the same index are said to be **coindexed** with each other
- Definition: NPs with the same index are said to **corefer** (i.e., refer to the same object in the outside world)

Binding Theory

Binding

- The relation between an antecedent and an anaphor/pronoun is a pretty rigid structural relation

    - *Ram$_i$ found himself$_i$ a blanket*

    - **Ram$_i$ found himself$_j$ a blanket*

    - *[The servant of Ram$_i$]$_j$ found himself$_i$ a blanket*

    - *[The servant of Ram$_i$]$_j$ found himself$_j$ a blanket*

Key Definitions

Domination

- Essentially the specification of a *tree* (very familiar to computer scientists!)
- Axioms of domination ($x <=_D y$ *means x dominates y*)
  - (a) $X <=_D X$
  - (b) if $X <=_D Y <=_D Z$ then $X <=_D Z$
  - (c) if $X <=_D Y <=_D X$ then $X = Y$
  - (d) if $X <=_D Z$ and $Y <=_D Z$ then either $X <=_D Y$ or $Y <=_D X$ (or both if $X=Y=Z$)

Immediate Domination and Exhaustive Domination

- **Immediate Domination**: Direct Parent Child relation

- **Exhaustive Domination**: Node A exhaustively dominates a *set* of nodes {B, C, …, D}, if it immediately dominates all the members of the set and *and* there is no node G immediately dominated by A that is not a member of this set.

Constituency

- **Constituent**: A set of nodes exhaustively dominated by a *single* node

- **Constituent-of**: B is a constituent of A iff A dominates B

- **Immediate-constituent-of**: B is an immediate-constituent-of A iff A immediately dominates B

Precedence ("said first" relation)

- S→NP VP
- S dominates NP and VP
- {NP VP} forms a constituent
- But NP *precedes* VP
- Definition: Node A **precedes** node B iff A is to the left of B and neither A dominates B nor B dominates A *and* every node dominating A either appears to the left of B or dominates B

No crossing branches constraint

- If one node X precedes another node Y then X and all nodes dominated by X must precede Y and all nodes dominated by Y

Axioms of Precedence

- Lets denote precedes by the symbol ~
- (a) If X ~ Y then NOT(Y ~ X)
- (b) If X~Y~Z then X~Z
- (c) If X~Y or Y~X then NOT(X $<=_D$ Y) and NOT(Y $<=_D$ X)
- (d) X~Y iff for all terminals U, V, X $<=_D$ U and Y $<=_D$ V jointly imply U~V

No crossing of branch; no discontinuous constituent

Fundamental concept: *c-command*

- (informal): A node **c-commands** its sisters and all the daughters (and granddaughters and great-granddaughters etc.) of its sisters.
- (formal): Node A **c-commands** node B if every branching node dominating A also dominates B, and neither A nor B dominate each other.

Example

M

O

N

A

B

C

D

E

M

G

H

I

J
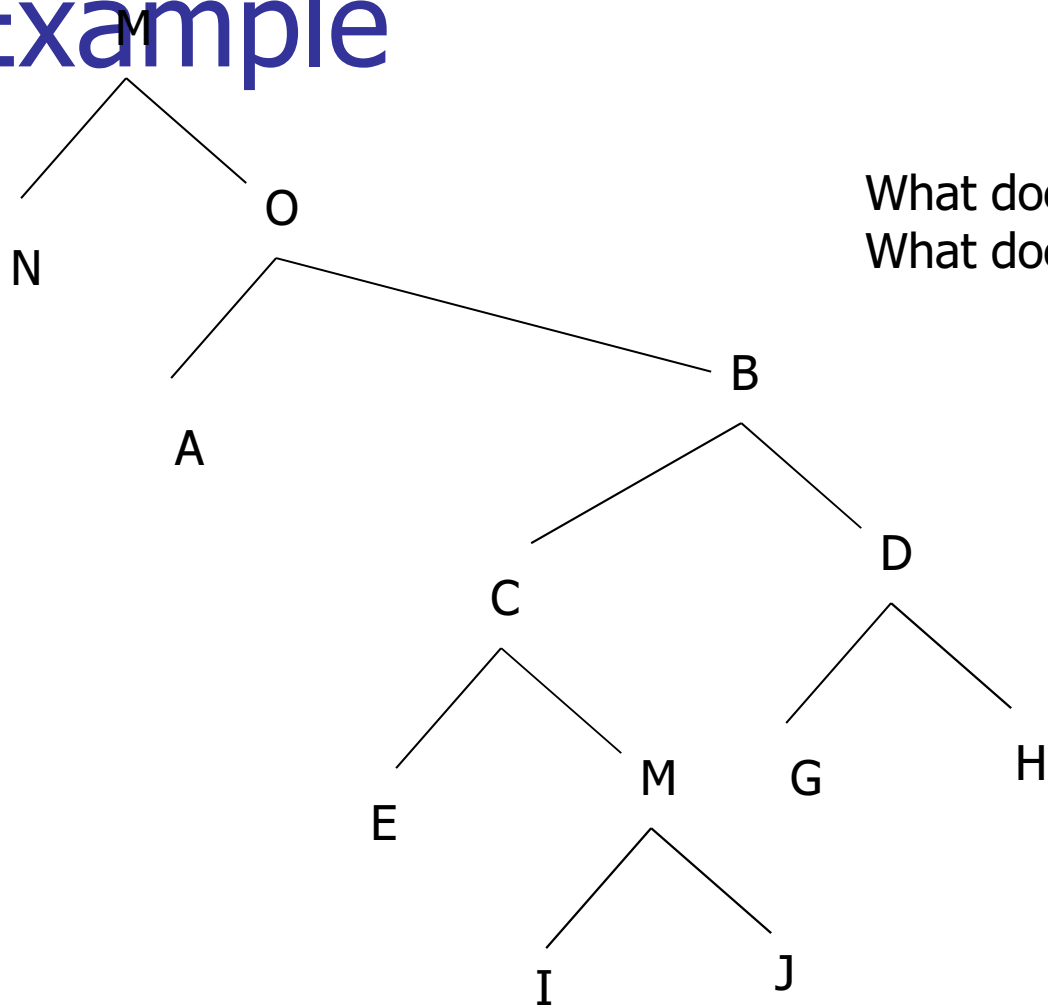
What does A c-command?
What does G c-command?

Example

M

O

N

B

A
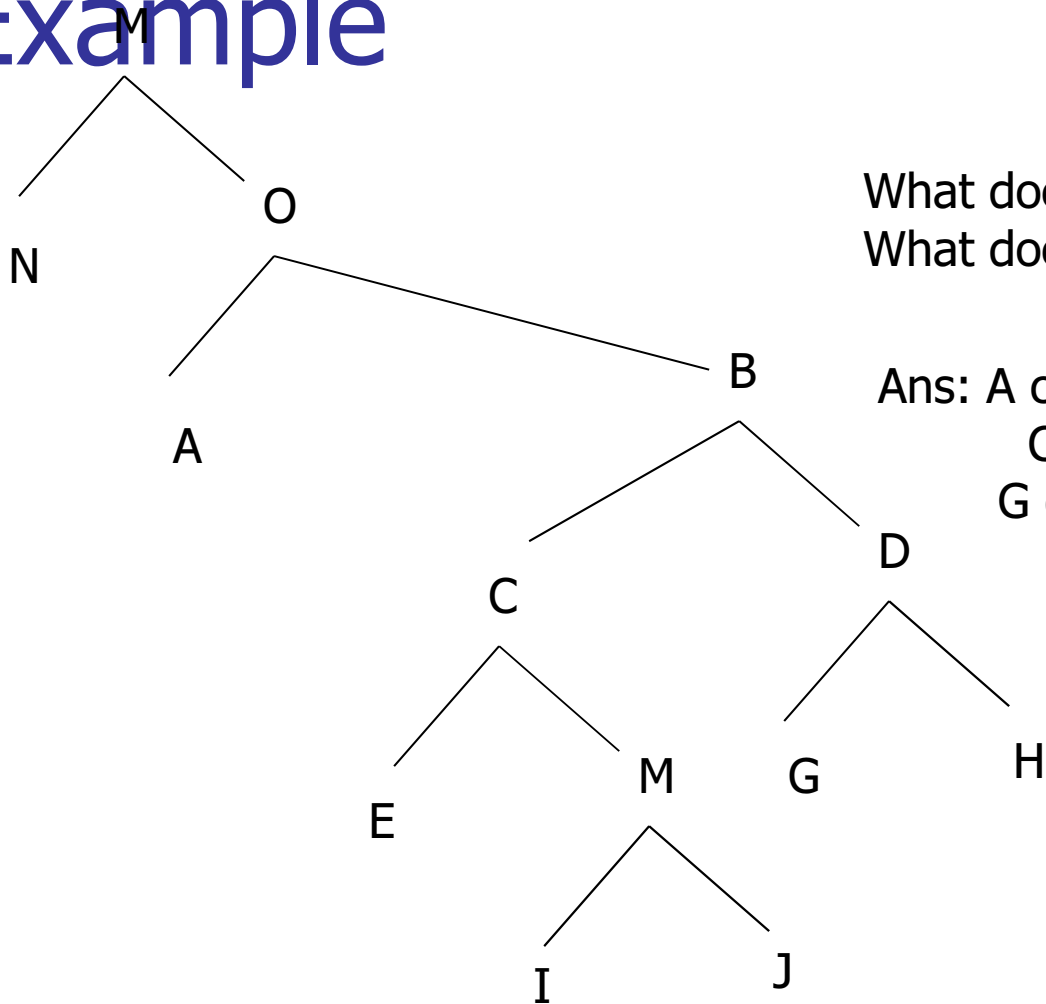
What does A c-command?
What does G c-command?

Ans: A c-commands B and
C,D,E,F,G,H,I,J
G c-commands only H

C

D

E

M

G

H

I

J

Symmetric C-command and Asymmetric c-command

- A symmetrically c-commands B, if A c-commands B and B c-commands A

- A asymmetrically c-commands B, if A c-commands B and B does not c-command A

Exercise

S

T
will

VP

NP₁

V
buy

NP₂

PP

D₁
The

AP

O

N₁
boy

N₂
apples

P
from

NP₃

D₃
the

N₃
grocer

little
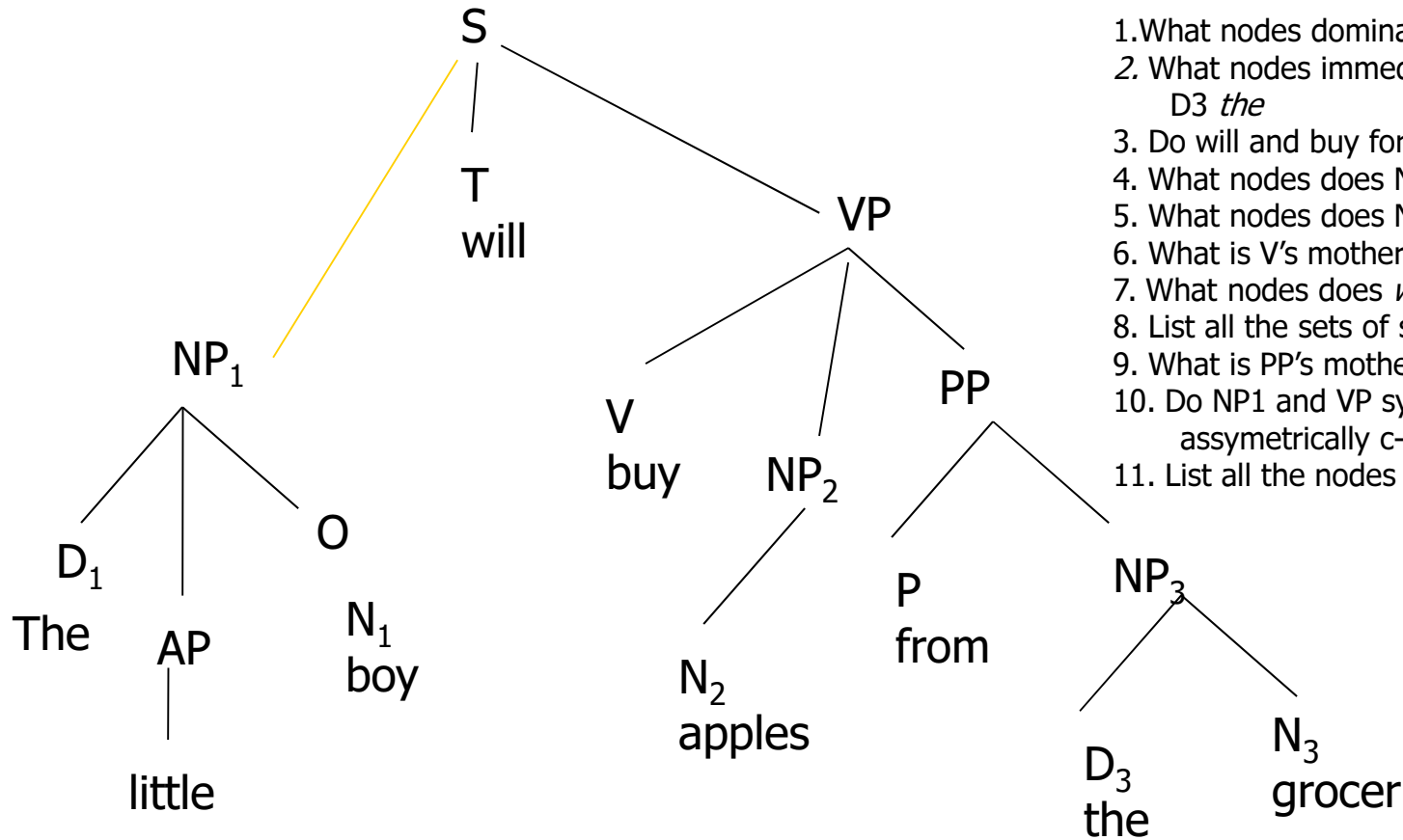
1. What nodes dominate *grocer*
2. What nodes immediately dominate D3 *the*
3. Do will and buy form a constituent?
4. What nodes does N₁ *boy* c-command?
5. What nodes does NP1 c-command?
6. What is V's mother?
7. What nodes does *will* precede?
8. List all the sets of sisters in the tree.
9. What is PP's mother?
10. Do NP1 and VP symmetrically or assymetrically c-command one another?
11. List all the nodes c-commanded by V

12. What is the subject of the sentence?
13. What is the object of the sentence?
14. What is the object of the preposition?
15. Is NP₃ a constituent of VP?

16. What node(s) in NP₃ an immediate constituent of ?
17. What node(s) does VP exhaustively dominate?
18. What is the root node?
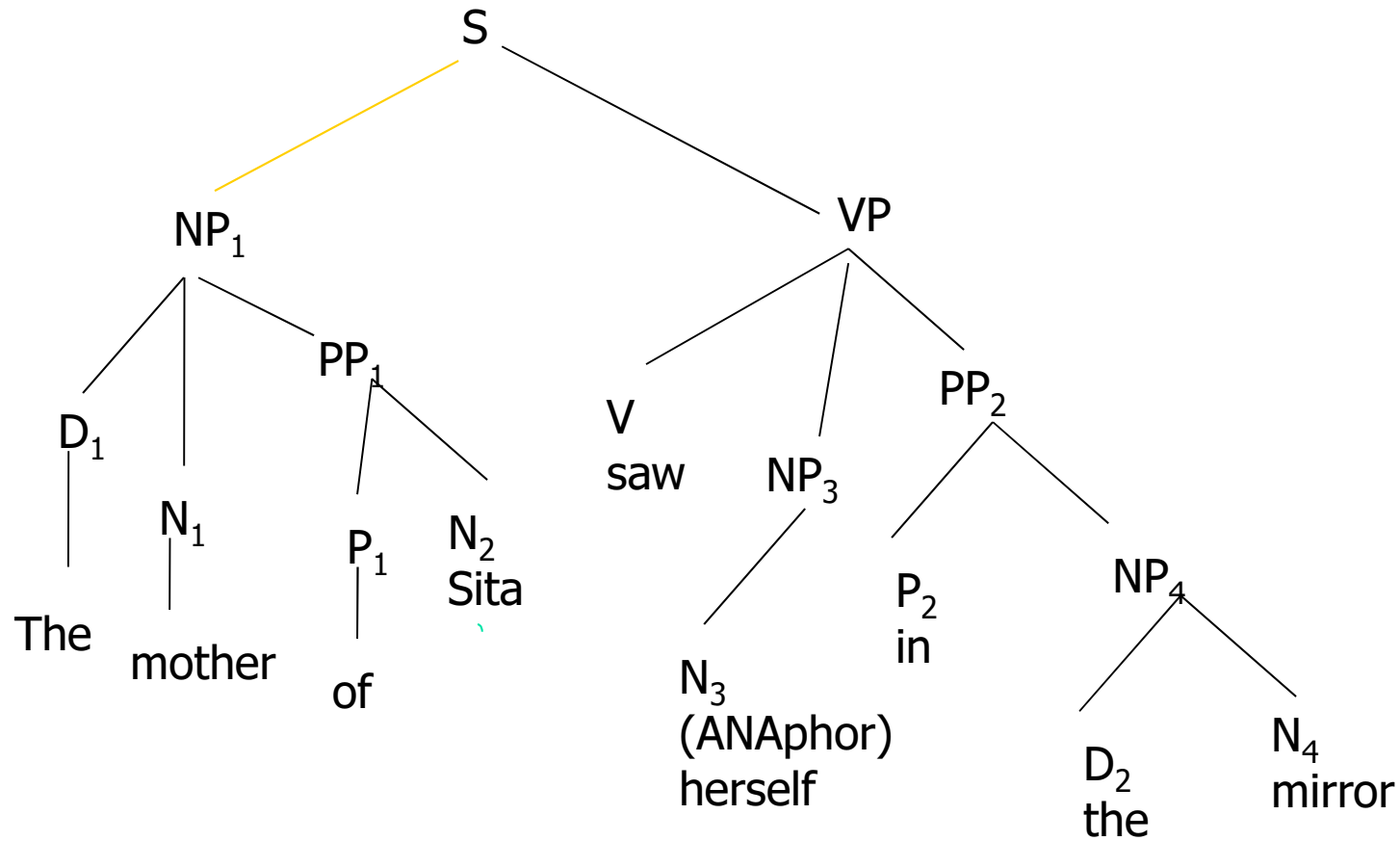19. List all the terminal nodes.
20. What immediately precedes *grocer?*

Correctness and incorrectness of binding

- *Sita$_p$ saw herself$_p$ in the mirror.*
- *[The mother of Sita$_q$]$_p$ saw herself$_p$ in the mirror.*
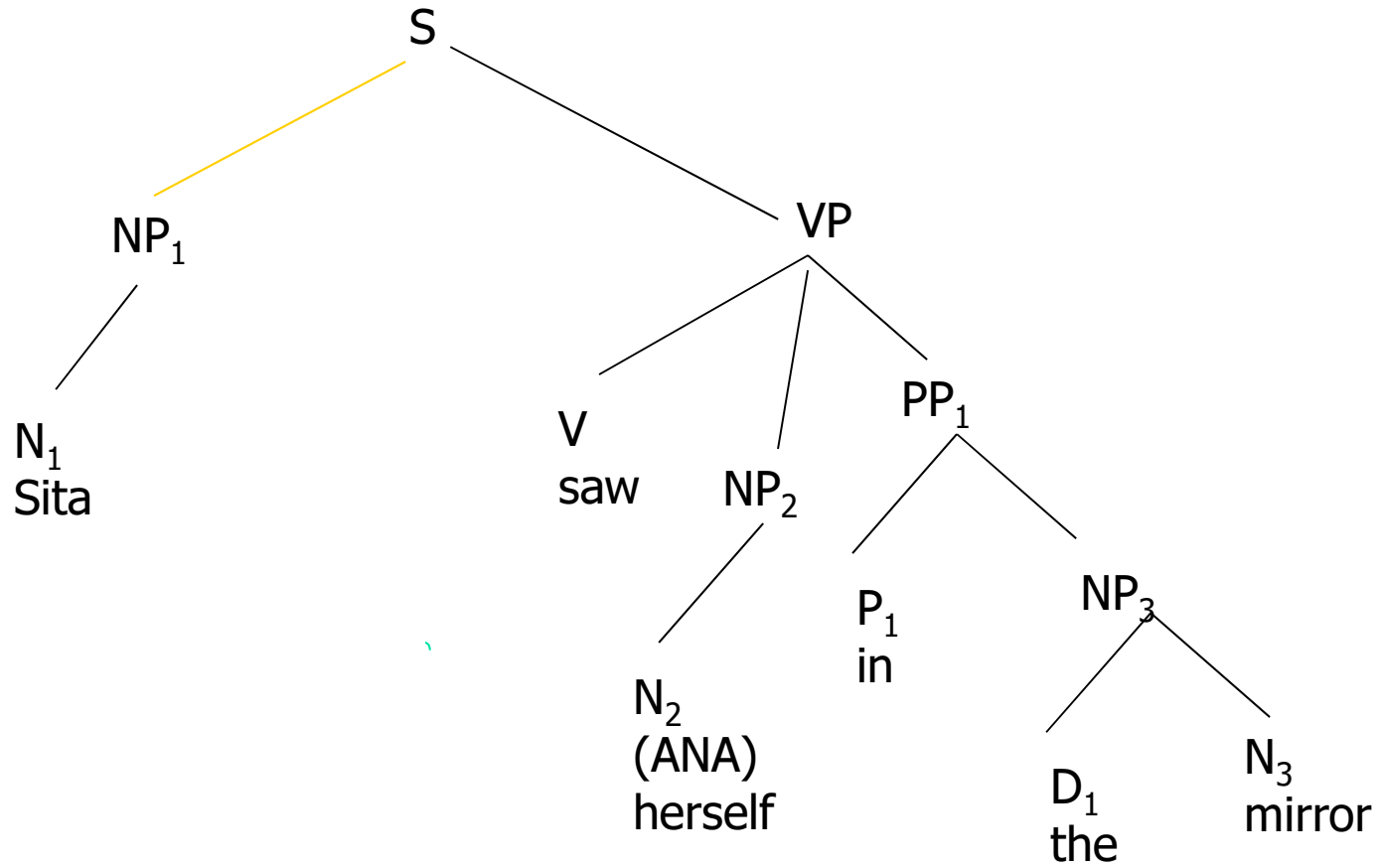- **[The mother of Sita$_q$]$_p$ saw herself$_q$ in the mirror.*

From the tree
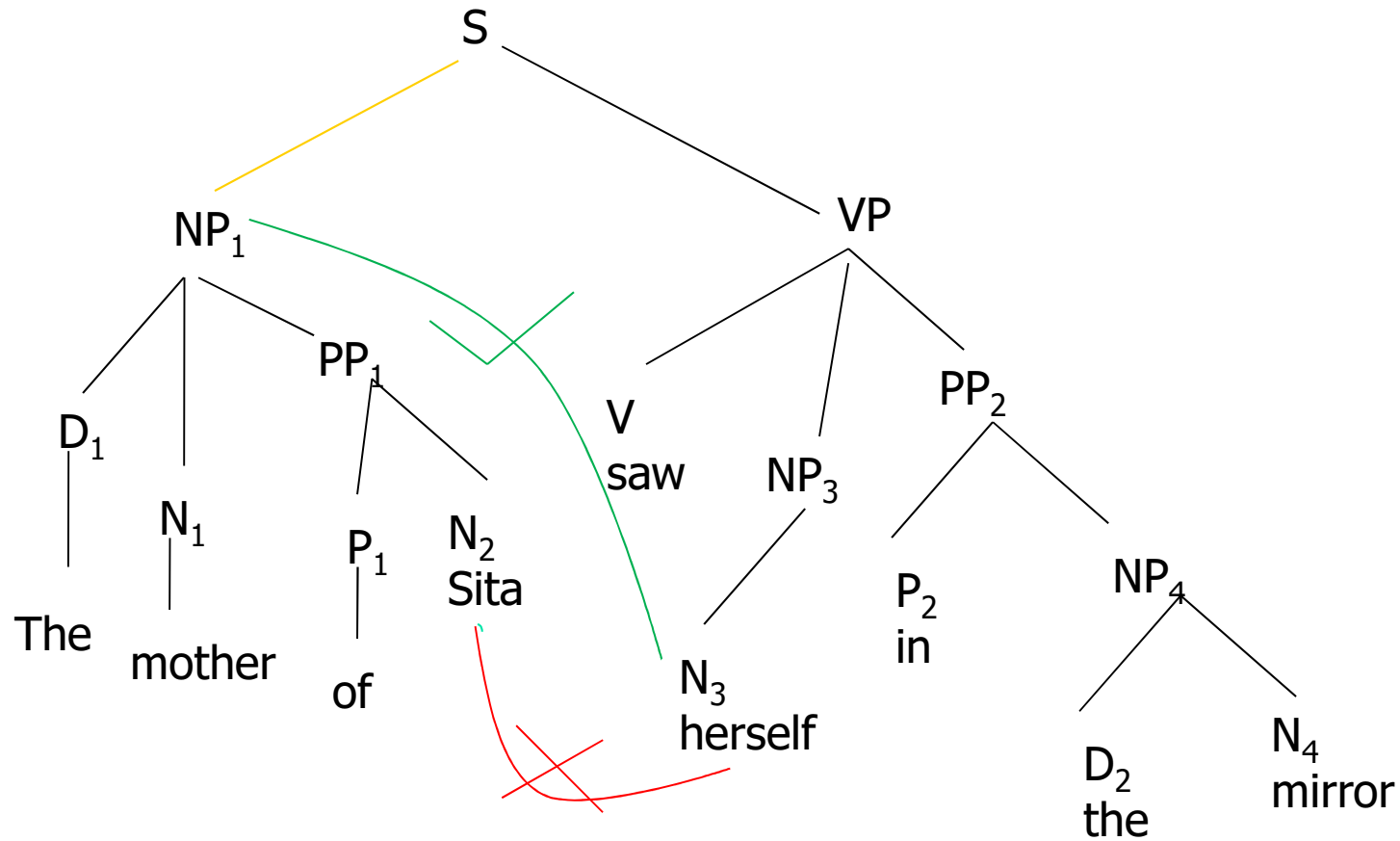
Case A

Case A Observations

- NP_{mother} -> herself ✓

- NP_{sita} -> herself ✗

- NP_{mother} -> her (meaning Sita) ✓

Case B

S
├── NP$_1$
│   └── N$_1$
│   Sita
└── VP
    ├── V
    │   saw
    ├── NP$_2$
    │   └── N$_2$
    │   (ANA)
    │   herself
    └── PP$_1$
        ├── P$_1$
        │   in
        └── NP$_3$
            ├── D$_1$
            │   the
            └── N$_3$
                mirror

Case B Observations

- NP_{sita} -> herself ✓

- NP_{sita} -> her ✗

Rules

- **Positive Rule of Binding for Anaphor**
  - Anaphor can be bound only to its c-commanding and preceding NP

- **Negative Rule of Binding for Pronoun**
  - Pronoun cannot be bound to a c-commanding NP

Definition of binding

- A binds B if
  - A c-commands B, and
  - A and B are coindexed
- Why is the following wrong?
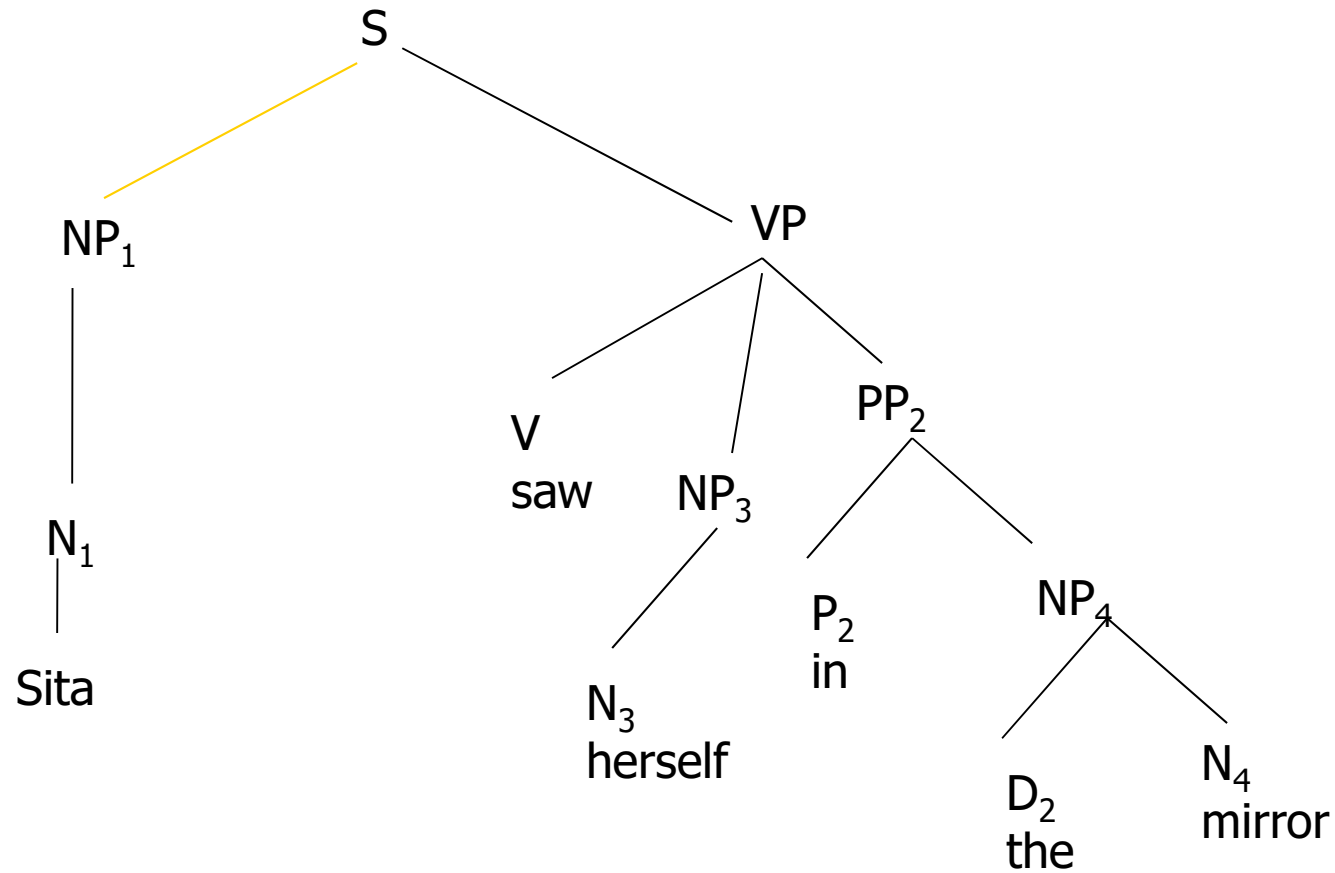 - *herself saw Sita in the mirror

Binding domain

- The syntactic space in which the anaphor must find its antecedent is called a *binding domain*.
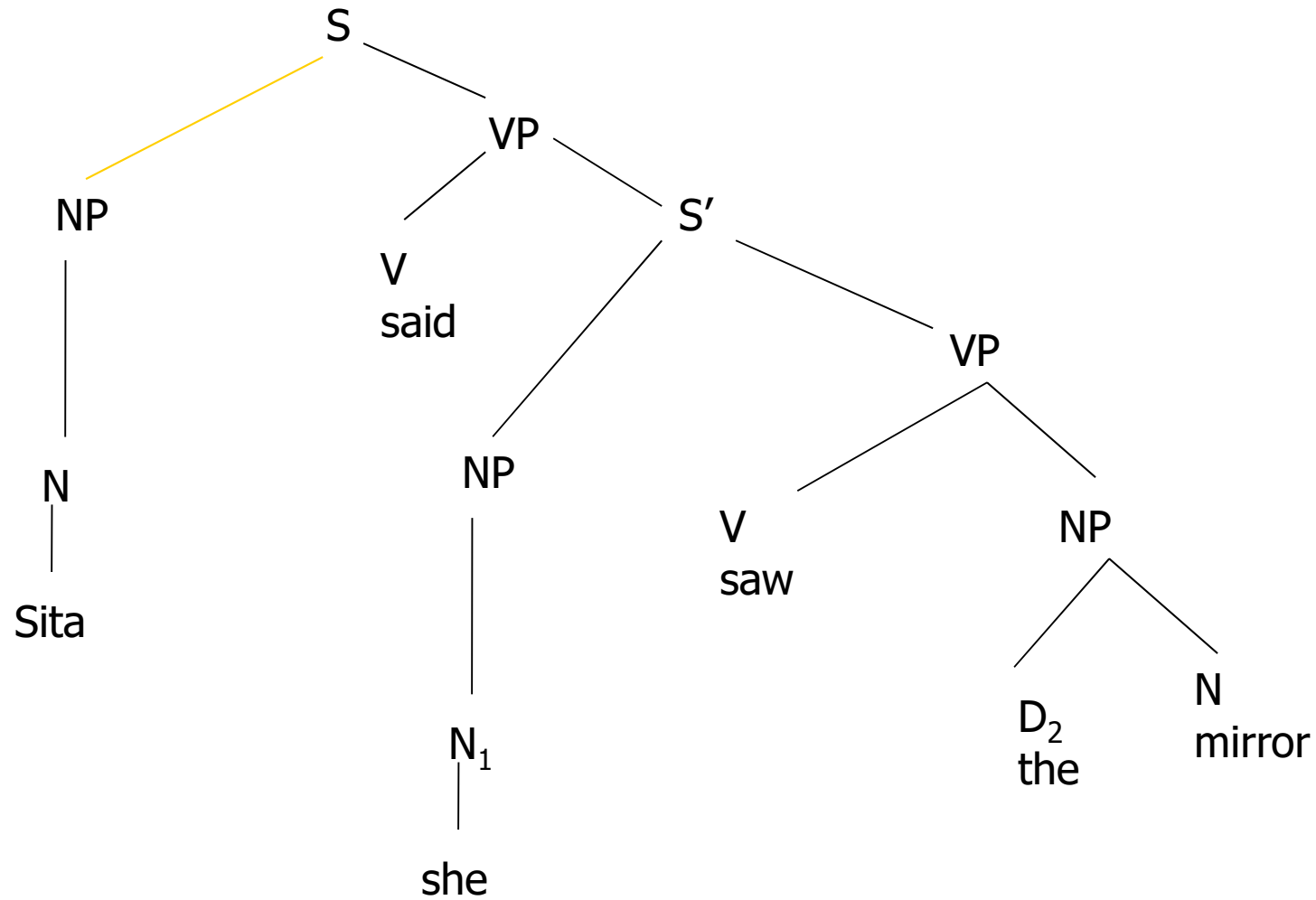
- Usually the binding domain is the clause.

Significance of binding domain

- *Sita saw herself in the mirror*
- *Sita said that she saw the mirror*
- **Sita said that herself saw the mirror*
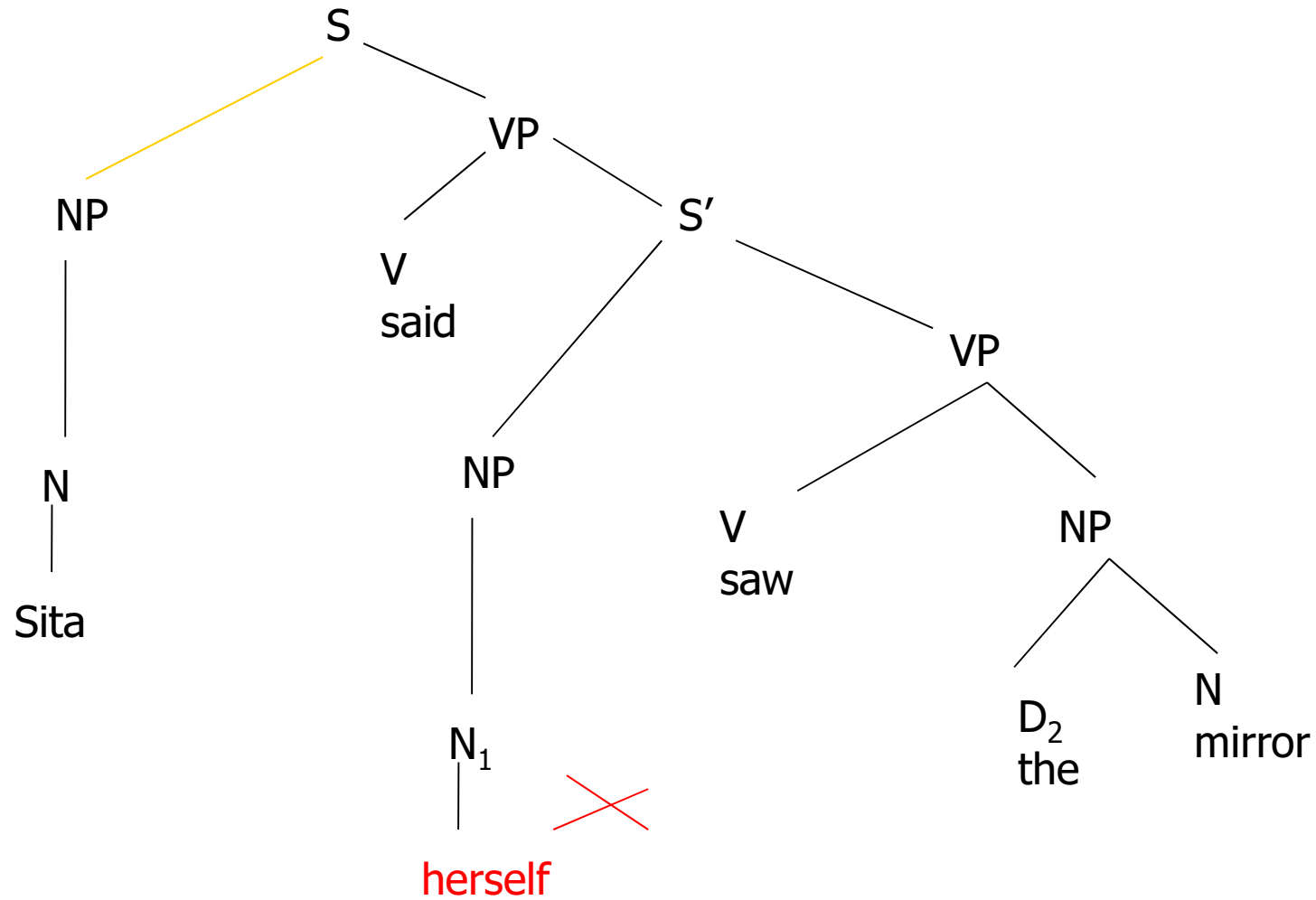- *Sita said that she saw herself in the mirror*

From the tree
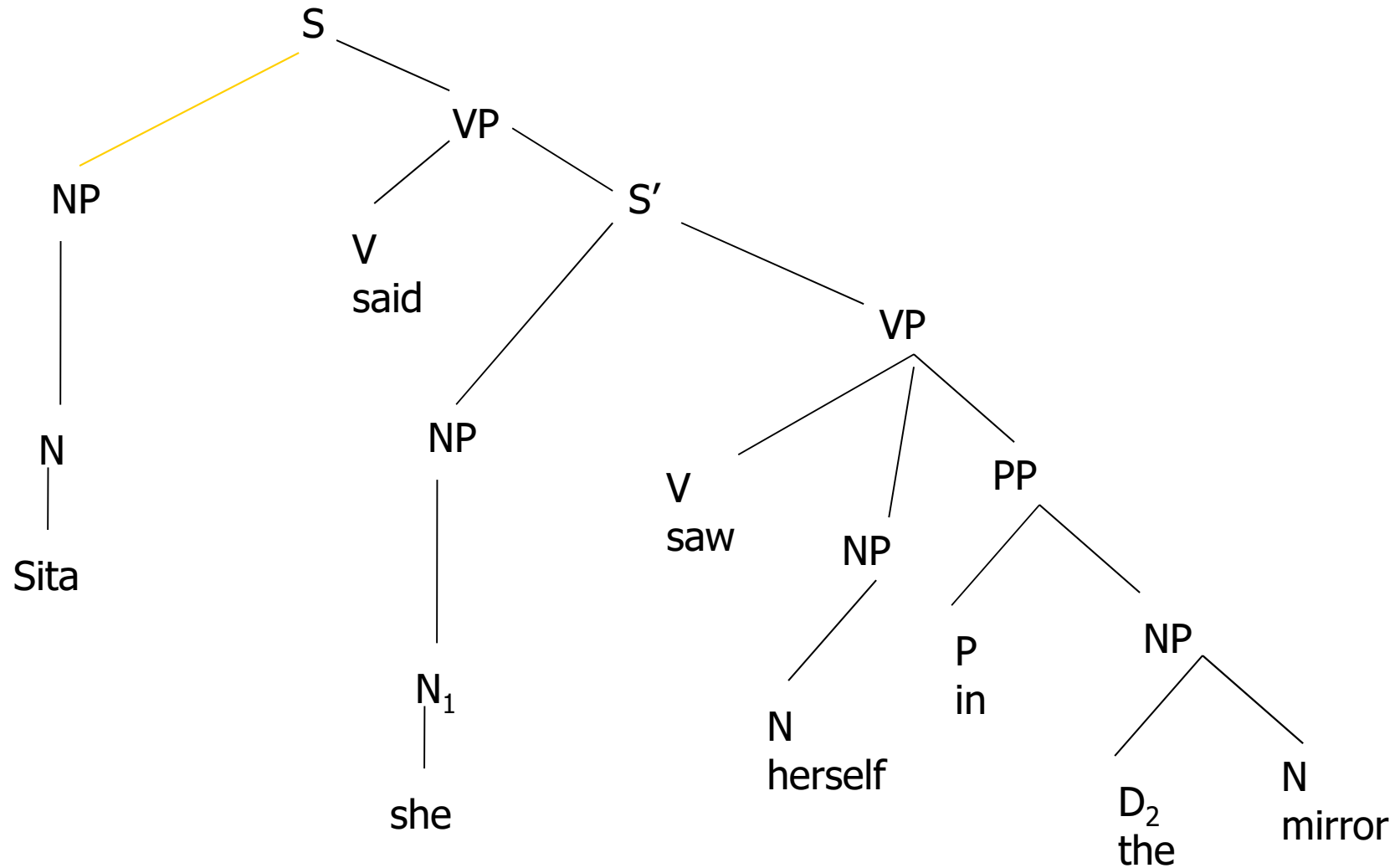
From the tree

From the tree

S
- NP
  - N
    - Sita
- VP
  - V
    - said
  - S′
    - NP
      - N₁
        - herself
    - VP
      - V
        - saw
      - NP
        - D₂
          - the
        - N
          - mirror

From the tree

Binding principle A

- An anaphor must be bound in its binding domain

From the tree

S
NP₁
VP
N₁
Sita_p
V
saw
NP₃
PP₂
N₃
her_p
P₂
in
NP₄
D₂
the
N₄
mirror

From the tree

S
├─ NP
│   └─ N
│   └─ Sita$_p$
└─ VP
    ├─ V
    │   └─ said
    └─ S'
        ├─ NP
        │   └─ N$_1$
        │   └─ She$_{p/q}$
        └─ VP
            ├─ V
            │   └─ saw
            ├─ NP
            │   └─ N
            │   └─ her$_r$
            └─ PP
                ├─ P
                │   └─ in
                └─ NP
                    ├─ D$_2$
                    │   └─ the
                    └─ N
                        └─ mirror
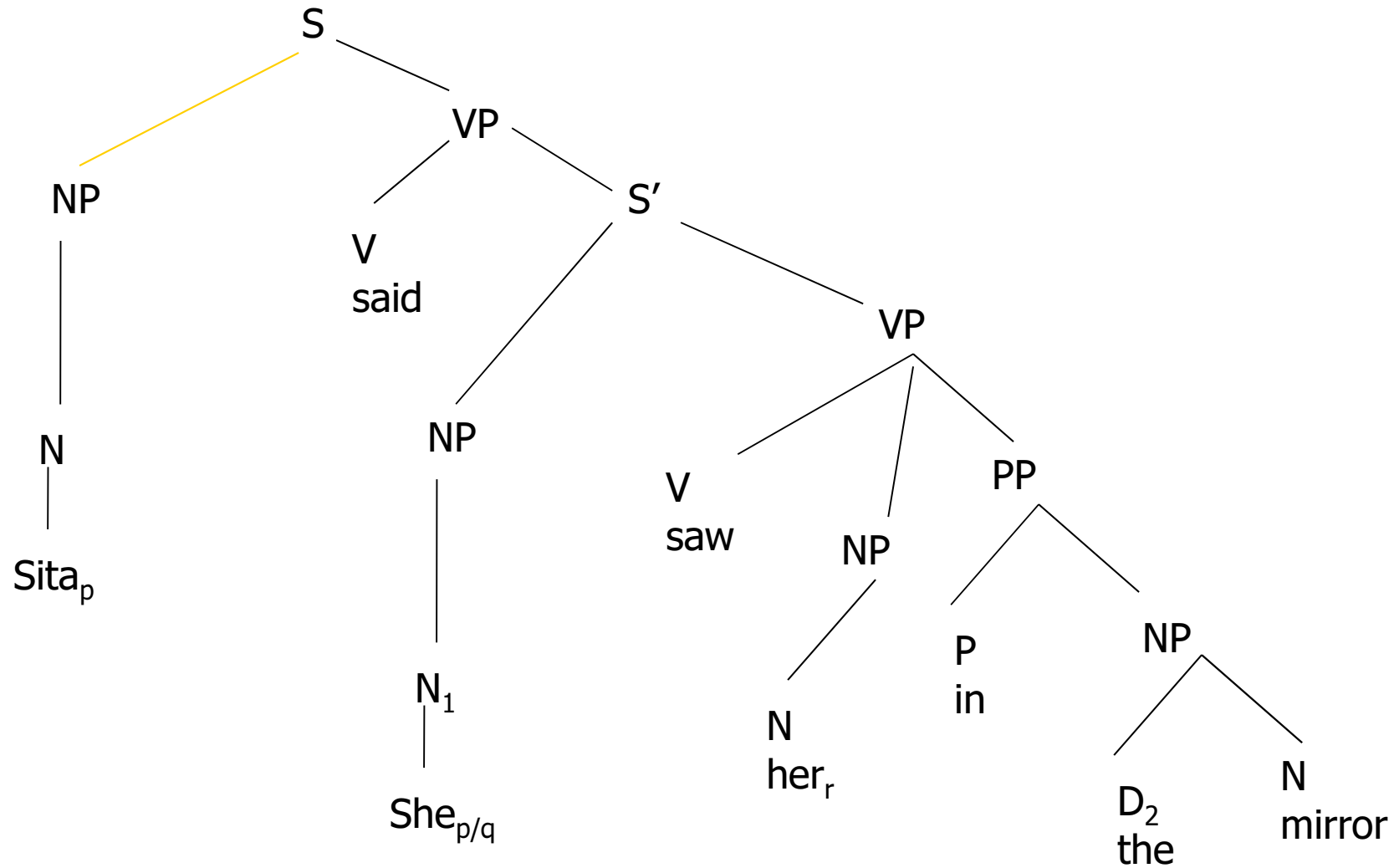
Binding principle B

- Definition: *Free-* not bound
- A pronoun must be free in its binding domain.

Binding principle C

- A R-expression (referring expression) must be free.