7. Transition-based parsing of CFGs

1 Embedding and acceptability patterns

The following collection of sentences provides a motivating "test set" for basic theories of human sentence processing.

(1) Left-branching structures
  a. Mary won
  b. Mary 's baby won
  c. Mary 's boss 's baby won

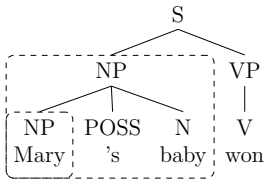
(2) Right-branching structures
  a. John met the boy
  b. John met the boy that saw the actor
  c. John met the boy that saw the actor that won the award

(3) Center-embedding structures
  a. the actor won
  b. the actor the boy met won
  → c. the actor the boy the baby saw met won

Here's a CFG generating all the sentences in (1), (2) and (3):

S → NP VP	N → baby, boy, actor, wife, boss
S → WHILE S S	NP → Mary, John
NP → NP POSS N	V → met, saw, won
NP → (D) N (PP) (SRC) (ORC)	D → the
VP → V (NP) (PP)	P → on, in, with
PP → P NP	THAT → that
SRC → THAT VP	POSS → 's
ORC → NP V	WHILE → while

Left-branching structure:



inside (Mary) (NP)
inside (Mary's baby) (NP)

the actor the boy the baby saw won

competence / performance

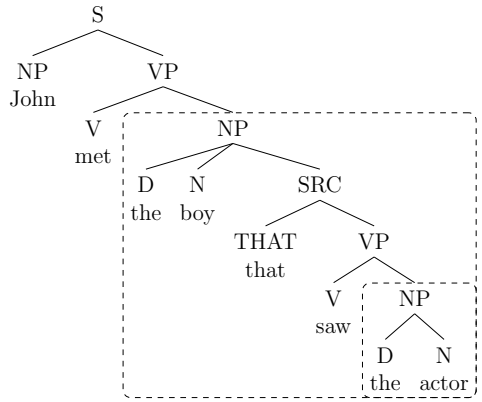
"performance difficulties"

"processing factors"

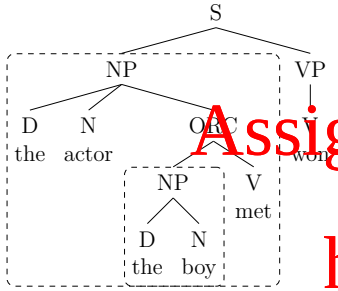
unacceptable

* who do you wonder who met?

Right-branching structure:



Center-embedding structure:



2 Transition-based parsing

A **parsing schema** defines, for any given grammar, a method for parsing/recognizing symbol sequences.
A transition-based parsing schema of the sort we will look at here has these components:

- a specification of what a **configuration** is;
- a specification of what a **starting configuration** is, for a given grammar and a given symbol-sequence;
- a specification of what a **goal configuration** is, for a given grammar;
- a specification of a **transition relation on configurations** (which I'll write as $\Longrightarrow$).

To give you a sense of what this means, let's start with the boringly-simple case of finite-state automata.

Parsing with an FSA

Starting configuration: $(A, x_1 \dots x_n)$
where A is a start state and $x_1 \dots x_n$ is the input

CONSUME step: $(A, x_i x_{i+1} \dots x_n) \Longrightarrow (B, x_{i+1} \dots x_n)$
where there is an arrow from A to B labeled with x_i

Goal configuration: (A, ϵ)
where A is a final state

3 CFG parsing schemas

In all of the following: A, B, etc. are placeholders for a nonterminal; x_1, x_2, etc. are placeholders for a terminal symbol; and $\mathbf{\Phi}$ is a placeholder for a sequence of nonterminals (in the form of a "stack").

We assume that the right-hand side of each CFG rule has *either* a single terminal symbol *or* a sequence of one-or-more nonterminal symbols.

$A \to BC$ $A \to B$
$A \to x$ $A \to BCD$

3.1 Bottom-up

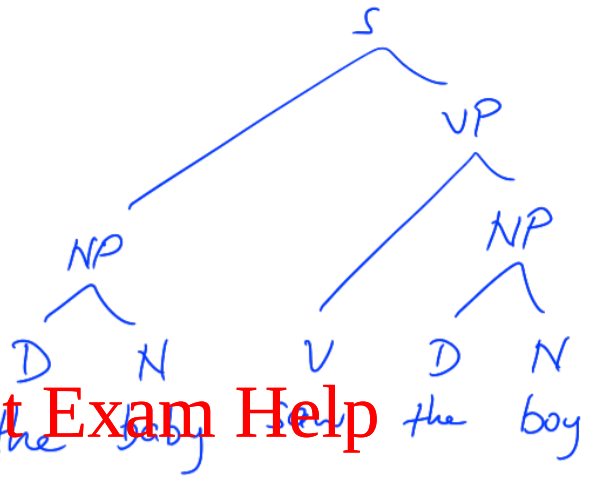
Bottom-up parsing schema

Starting configuration: $(\epsilon, x_1 \dots x_n)$
where $x_1 \dots x_n$ is the input

SHIFT step: $(\mathbf{\Phi}, x_i x_{i+1} \dots x_n) \Longrightarrow (\mathbf{\Phi} A, x_{i+1} \dots x_n)$
where there is a rule $A \to x_i$ in the grammar

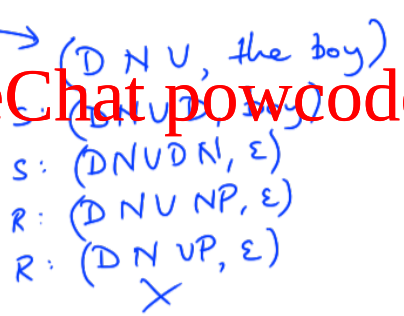
REDUCE step: $(\mathbf{\Phi} B_1 \dots B_m, x_i \dots x_n) \Longrightarrow (\mathbf{\Phi} A, x_i \dots x_n)$
where there is a rule $A \to B_1 \dots B_m$ in the grammar

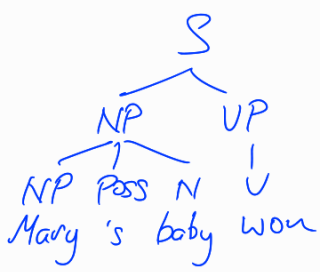
Goal configuration: (A, ϵ)
where A is one of the grammar's start symbols



Example:

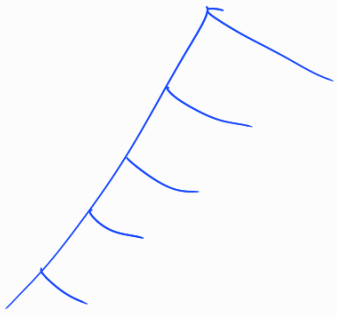
	Type of transition	Rule used	Configuration
0	—	—	(ϵ, the baby saw the boy)
1	SHIFT	D → the	(D, baby saw the boy)
2	SHIFT	N → baby	(D N, saw the boy)
3	REDUCE	NP → D N	(NP, saw the boy)
4	SHIFT	V → saw	(NP V, the boy)
5	SHIFT	D → the	(NP V D, boy)
6	SHIFT	N → boy	(NP V D N, ϵ)
7	REDUCE	NP → D N	(NP V NP, ϵ)
8	REDUCE	VP → V NP	(NP VP, ϵ)
9	REDUCE	S → NP VP	(S, ϵ)



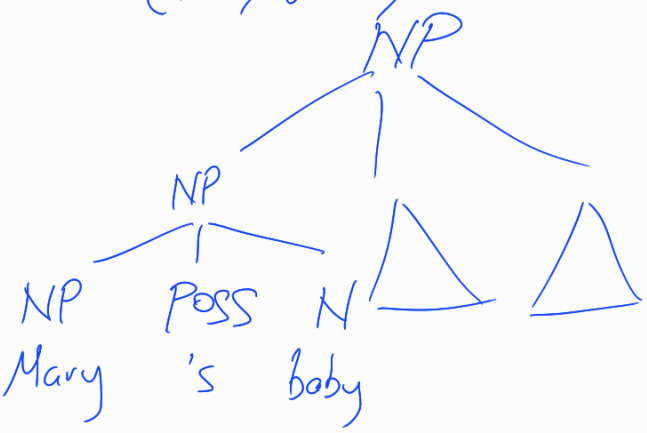


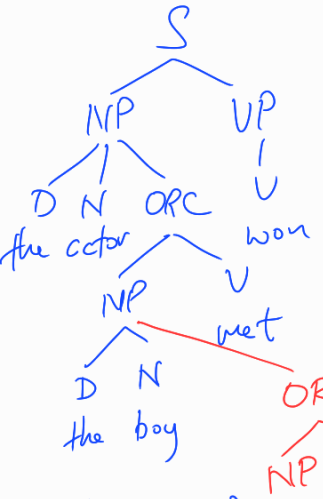
(ε, Mary 's baby won)
(NP, 's baby won)
(NP POSS, baby won)
(NP POSS N, won)
(NP, won)
(NP V, ε)
(NP VP, ε)
(S, ε)

BU
LE
3,3



(ε, Mary 's boss 's baby won)
(NP, 's boss 's baby won)
(NP POSS, boss 's baby won)
(NP POSS N, 's baby won)
(NP, 's baby won)
(NP POSS, baby won)
(NP POSS N, won)
(NP, won)





(ε, 0)
(D, 1)
(D N, 2)
(D N D N, 4)
(D N NP, 4)
(D N NP V, 5)
(D N ORC, 5)
⋮
(S, 6)

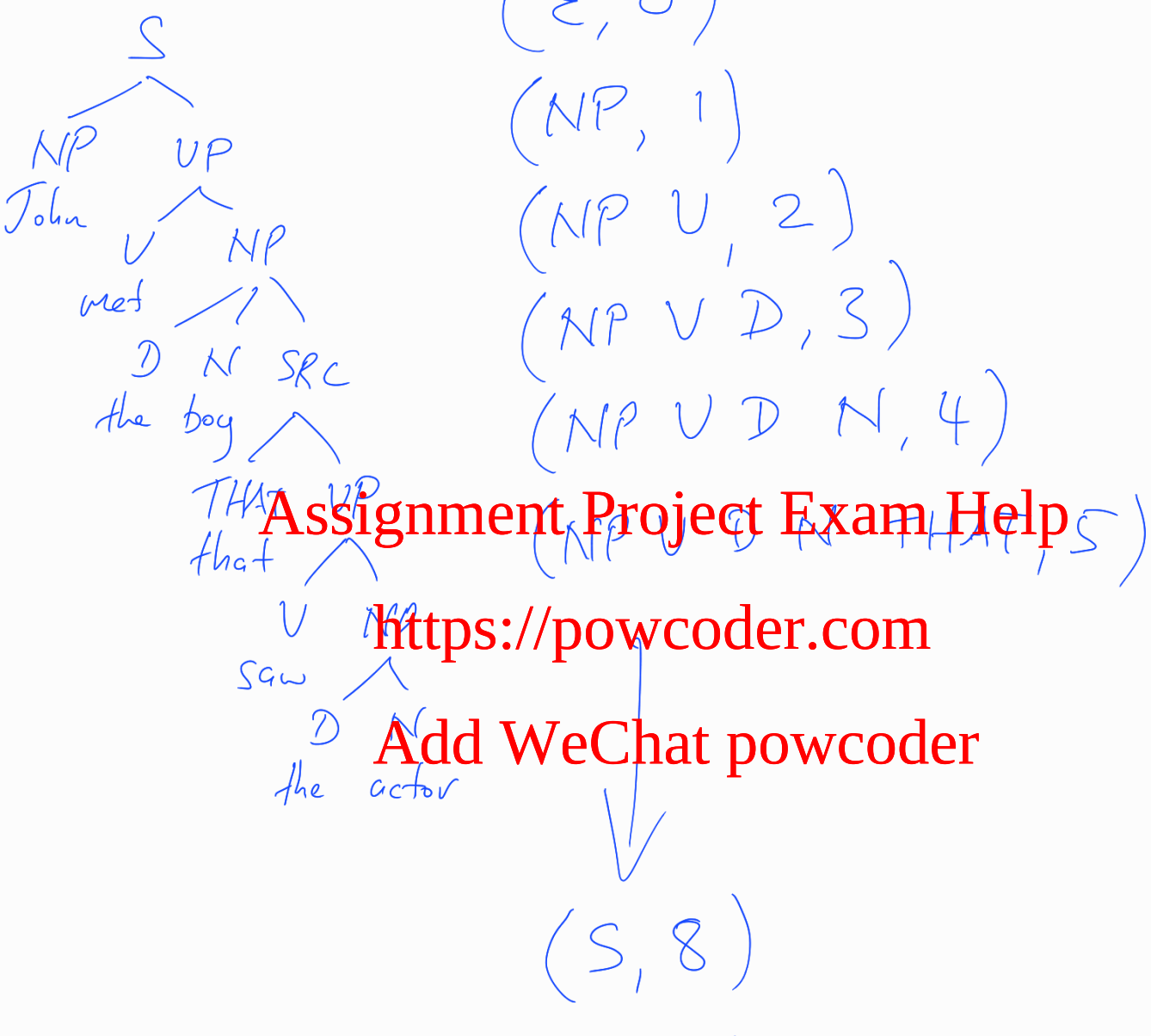
the actor that the boy met won the baby saw
0 1 2 3 4 5 6 7 8 9

(ε, 0)
(D N, 2)
(D N D N, 4)
(D N D N D, 5)
(D N D N D N, 6)
(D N D N NP, 6)
(D N D N NP V, 7)
(D N D N ORC, 7)
(D N NP, 7)

BU
CE
4,6

BU
RE
8,12

John 0 met 1 the 2 boy 3 that 4 saw 5 the 6 actor 7 8



(ε, 0)

(NP, 1)

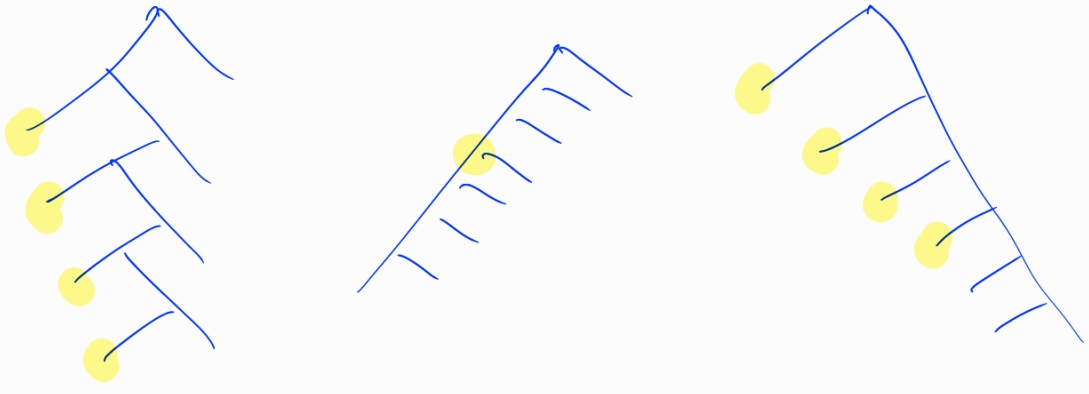
(NP V, 2)

(NP V D, 3)

(NP V D N, 4)

(NP V D N THAT, 5)

(S, 8)



3.2 Top-down

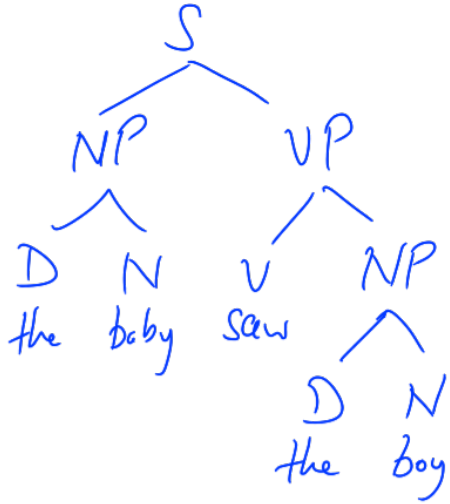
Top-down parsing schema

Starting configuration: $(A, x_1 \dots x_n)$
where A is one of the grammar's start symbols and $x_1 \dots x_n$ is the input

PREDICT step: $(A\mathbf{\Phi}, x_i \dots x_n) \Longrightarrow (B_1 \dots B_m\mathbf{\Phi}, x_i \dots x_n)$
where there is a rule $A \to B_1 \dots B_m$ in the grammar

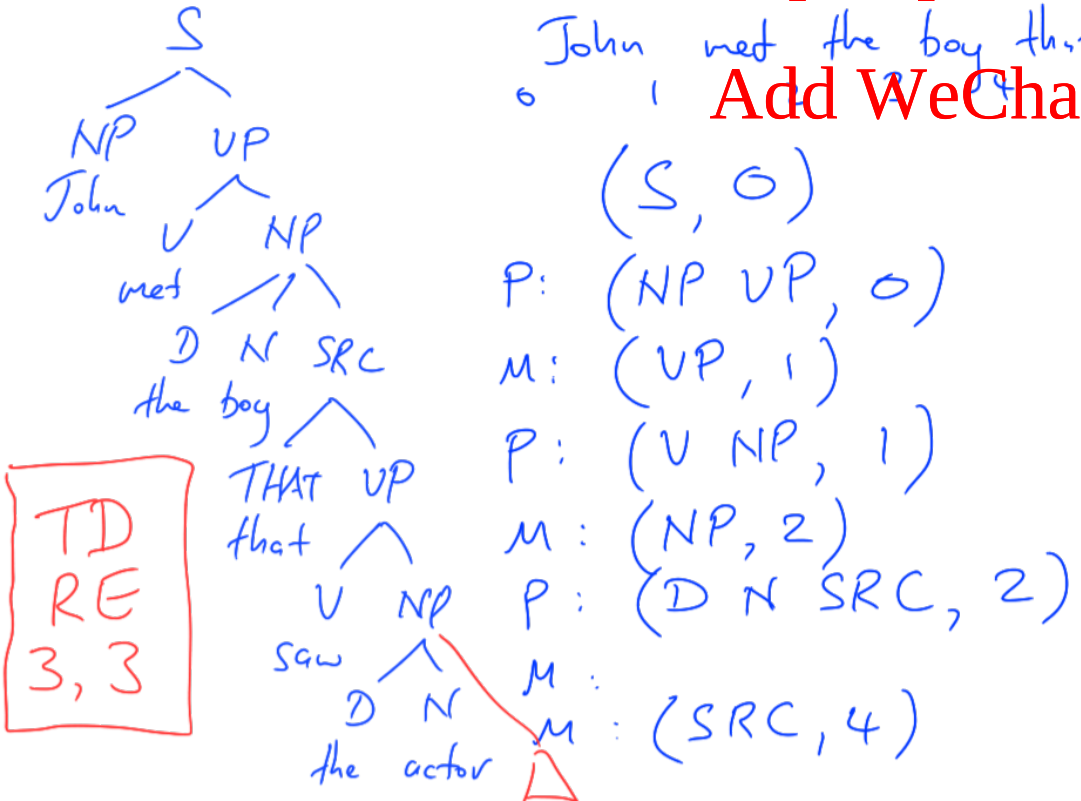
MATCH step: $(A\mathbf{\Phi}, x_i x_{i+1} \dots x_n) \Longrightarrow (\mathbf{\Phi}, x_{i+1} \dots x_n)$
where there is a rule $A \to x_i$ in the grammar

Goal configuration: (ϵ, ϵ)

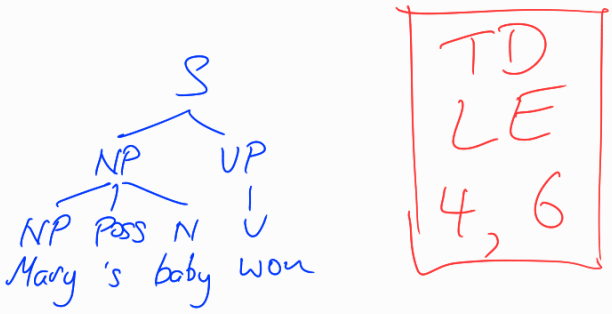


Example:

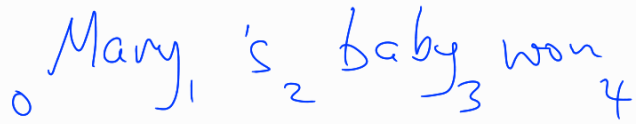
	Type of transition	Rule used	Configuration
0	—	—	(S, the baby saw the boy)
1	PREDICT	S → NP VP	(NP VP, the baby saw the boy)
2	PREDICT	NP → D N	(D N VP, the baby saw the boy)
3	MATCH	D → the	(N VP, baby saw the boy)
4	MATCH	N → baby	(VP, saw the boy)
5	PREDICT	VP → V NP	(V NP, saw the boy)
6	MATCH	V → saw	(NP, the boy)
7	PREDICT	NP → D N	(D N, the boy)
8	MATCH	D → the	(N, boy)
9	MATCH	N → boy	(ϵ, ϵ)



	LE	RE	CE
HUMANS	no incr.	no incr.	increase
BU	3→3 no increase	8→12 increase	4→6 increase
TD	4→6 incr.	3→3 no incr.	
LC			



TD
LE
4, 6



(S, 0)

(NP VP, 0)

M: (VP, 1)

X

P: (NP POSS N VP, 0)

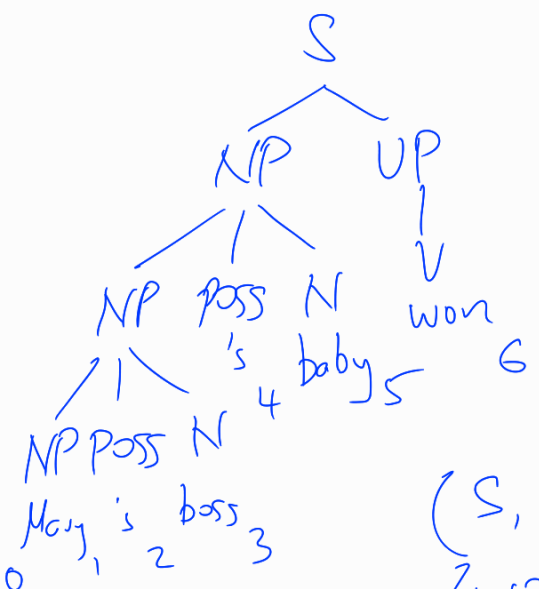
M

M

M: (VP, 3)

P: (V, 3)

M: (ε, 4)

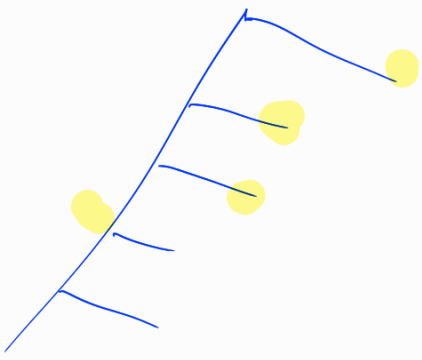


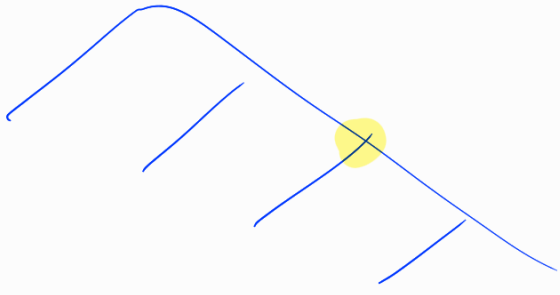
(S, 0)

(NP VP, 0)

(NP POSS N VP, 0)

(NP POSS N VP, 0)





3.3 Left-corner

Left-corner parsing schema

Starting configuration: $(\overline{A}, x_1 \dots x_n)$
where A is one of the grammar's start symbols and $x_1 \dots x_n$ is the input

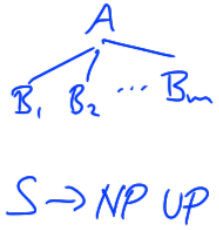
SHIFT step: $(\mathbf{\Phi}, x_i x_{i+1} \dots x_n) \Longrightarrow (A\mathbf{\Phi}, x_{i+1} \dots x_n)$
where there is a rule $A \to x_i$ in the grammar

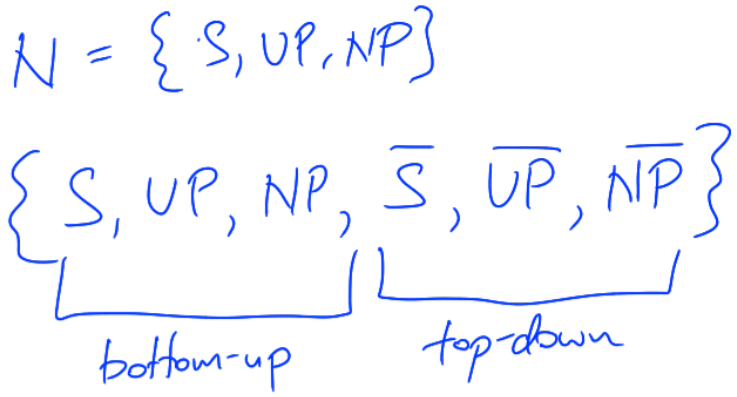
MATCH step: $(\overline{A}\mathbf{\Phi}, x_i x_{i+1} \dots x_n) \Longrightarrow (\mathbf{\Phi}, x_{i+1} \dots x_n)$
where there is a rule $A \to x_i$ in the grammar

LC-PREDICT step: $(B_1\mathbf{\Phi}, x_i \dots x_n) \Longrightarrow (\overline{B_2} \dots \overline{B_m} A\mathbf{\Phi}, x_i \dots x_n)$
where there is a rule $A \to B_1 \dots B_m$ in the grammar

LC-CONNECT step: $(B_1\overline{A}\mathbf{\Phi}, x_i \dots x_n) \Longrightarrow (\overline{B_2} \dots \overline{B_m}\mathbf{\Phi}, x_i \dots x_n)$
where there is a rule $A \to B_1 \dots B_m$ in the grammar

Goal configuration: (ϵ, ϵ)





Example:

	Type of transition	Rule used	Configuration
0	—	—	($\overline{S}$, the baby saw the boy)
1	SHIFT	D → the	(D $\overline{S}$, baby saw the boy)
2	LC-PREDICT	NP → D N	($\overline{N}$ NP $\overline{S}$, baby saw the boy)
3	MATCH	N → baby	(NP $\overline{S}$, saw the boy)
4	LC-CONNECT	S → NP VP	($\overline{VP}$, saw the boy)
5	SHIFT	V → saw	(V $\overline{VP}$, the boy)
6	LC-CONNECT	VP → V NP	($\overline{NP}$, the boy)
7	SHIFT	D → the	(D $\overline{NP}$, boy)
8	LC-CONNECT	NP → D N	($\overline{N}$, boy)
9	MATCH	N → boy	(ϵ, ϵ)