

Nov 8

COMPSCI 3331

Fall 2022

What's next?

- ▶ Assignment 3: available now.
- ▶ Quiz 4 marks available now.
- ▶ Quiz 5 tomorrow - **Lectures 8 and 9.**
- ▶ Midterm marks: soon.
- ▶ Assignment 2: being marked.

~~CYK~~ - Useless Symbols

CNF

▷ is generating
↑

$S \rightarrow aaBaA \mid aE \mid \underline{caDbb}$

$A \rightarrow E \mid ABaa \mid ab$

$B \rightarrow aB \mid \varepsilon$

~~C~~ $\rightarrow CC \mid a$ C is unreachable.

$D \rightarrow aba$

~~E~~ $\rightarrow \underline{EEF \mid aEaa}$ E is the only non-generating one \Rightarrow it is useless.

~~F~~ $\rightarrow aF \mid a\underline{G} \mid ababb$

~~G~~ $\rightarrow GG \mid aGa \mid \varepsilon$ Since F is reachable, G is reachable.

↓
 F and G become useless after we remove
 E from the language.

Useful:

1. reachable

2. generating

generating: $ABCD FGS$

variable that could get letters in one step. i.e. goto sequence of letter

reachable: $ABDEFGS$ / including ε .

generating - immediately

- goto another generating stage.

a variable that is not generate would cause infinite loop. So it could be removed directly.

CYK Algorithm $O(n^3)$ non-terminal

= Find a word is in the grammar or not.

Asmt. Q3. $S \rightarrow AB \mid CD$

$A \rightarrow AE \mid a$

$B \rightarrow EF \mid \#$

$C \rightarrow EC \mid a$

$D \rightarrow SG \mid \#$

$E \rightarrow a$

$F \rightarrow BG$

$G \rightarrow b$

e.g. $aaa\#b$

the box could
be empty, and
it is still possible
to generate the
answer.

once found S
is in $U(1)$, this algorithm
could be terminated.



△ if S is in the last cell, then the word is in grammar
but it does not tell you how the word is generated.

letter

non-terminal

the non-terminal
that could generate one of the
combination in $(ACE) \times (ACE)$.

i.e. $C \rightarrow EC$, so does A .

ACE ACE

$$U(3,1) = U(4,1) \times U(3,2) + U(5,1) \times U(3,3)$$

$$= AC \times ACE + ACE \times AC$$

$$= \{AC, AA, AE, CA, CC, CE, EA, EC, EE\}$$

$$U(2,2) = U(3,2) \times U(2,4) + U(4,2) \times U(2,3)$$

$$= ACE \times S + AC \times BD$$

$$= \{AS, CS, ES, AB, AD, CB, CD\}$$

$$U(1,3) = \{SG, AF, CF, EF\}$$

+len...

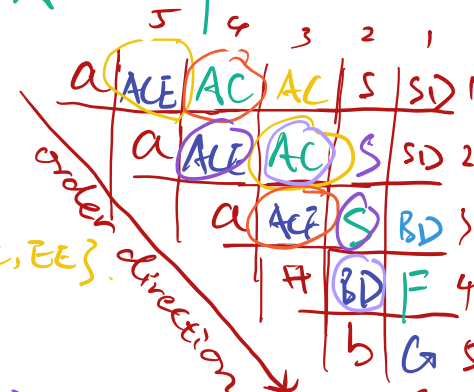
$$U(2,1) = U(5,1) \times U(2,2) + U(5,1) \times U(2,3) +$$

$$U(4,1) \times U(2,2) + U(4,1) \times U(2,4) +$$

$$U(3,1) \times U(2,3) + U(3,1) \times U(2,4)$$

$$U(1,2) = \dots$$

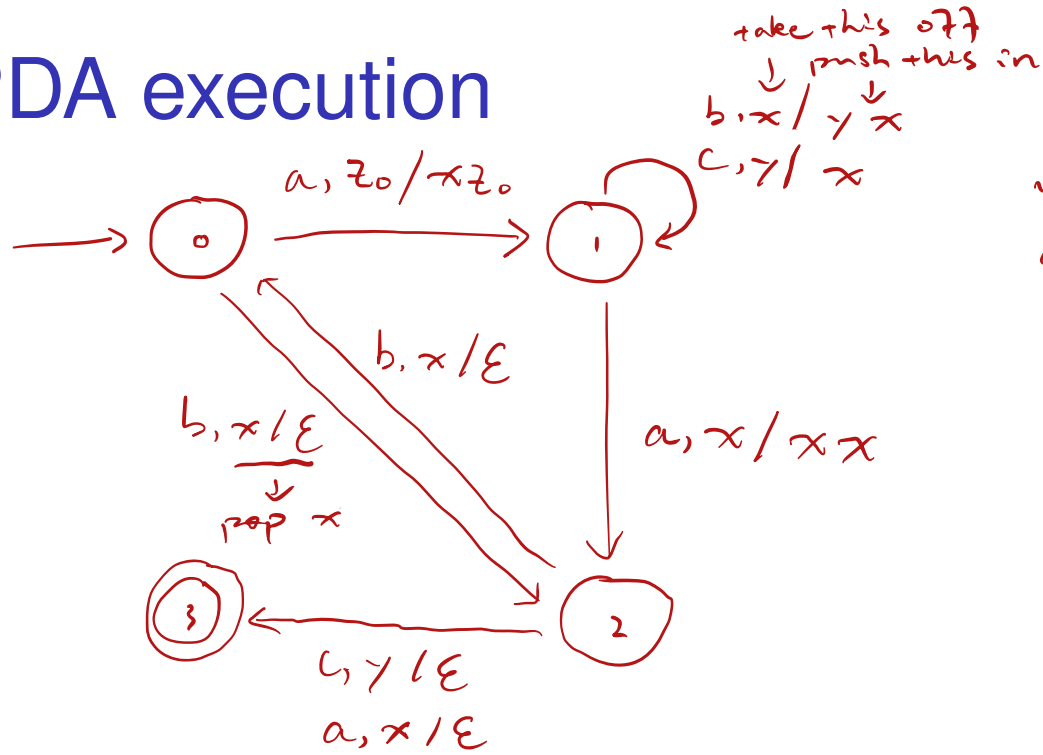
$$U(1,1) = \dots$$



cells are non-terminal that
can be derived
from subword

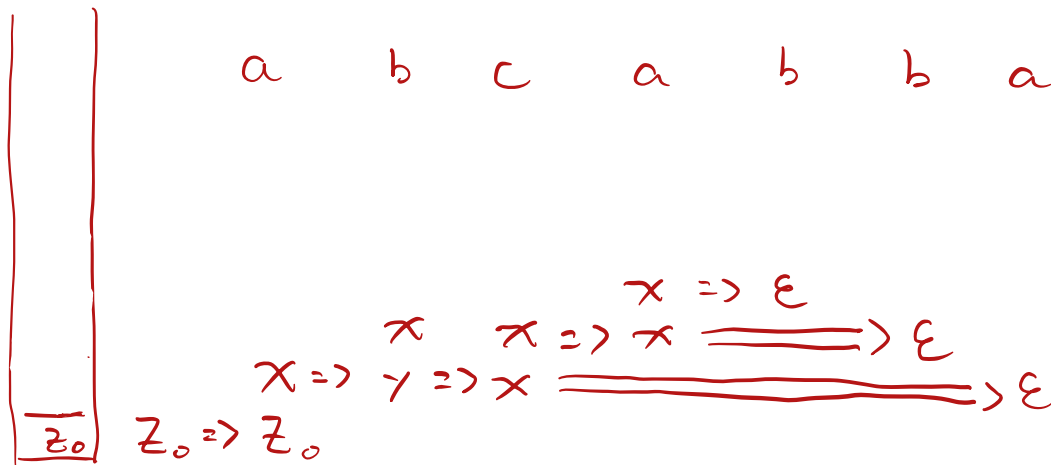
Base case: non-terminals that
generate subword
it is dynamic
programming.

PDA execution



what language does this PDA accepts?

random input: a b c a b b a

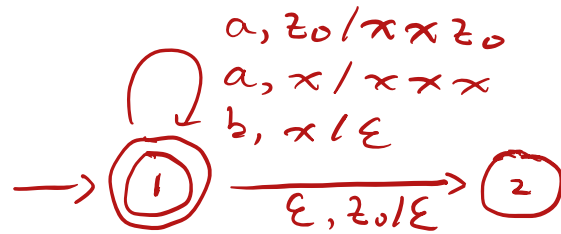


the word is accepted if the stack:

1. able to pop all elements (including z_0)
2. pops all elements except z_0 AND stops at a final state.

PDA's

$$L = \{a^n b^{2n} : n \geq 0\}$$



ends at 1 with z_0
at the end of the stack.

PDAs

- ▶ ★☆☆ $L_1 = \{a^n b^m : n, m \geq 0, n \neq m\}$
- ▶ ★☆☆ B , the language of balanced parentheses over $(,)$.
- ▶ ★★☆☆ $L_2 = \{w \in \{a, b\}^* : |w|_a = |w|_b\}$
- ▶ ★★☆☆ $L_3 = \{a^i b^j c^k : i = j \text{ or } i = k\}$.