CMSC 141 Automata and Language Theory

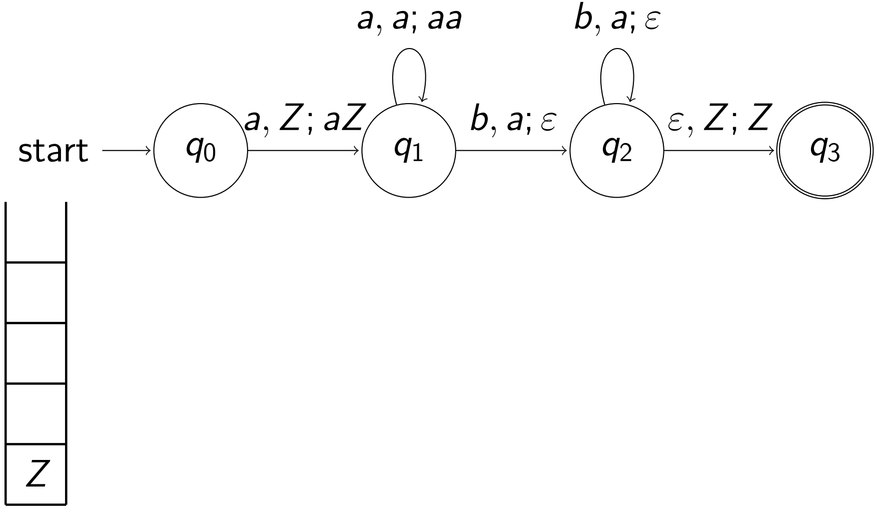
Context-Free Languages

Mark Froilan B. Tandoc

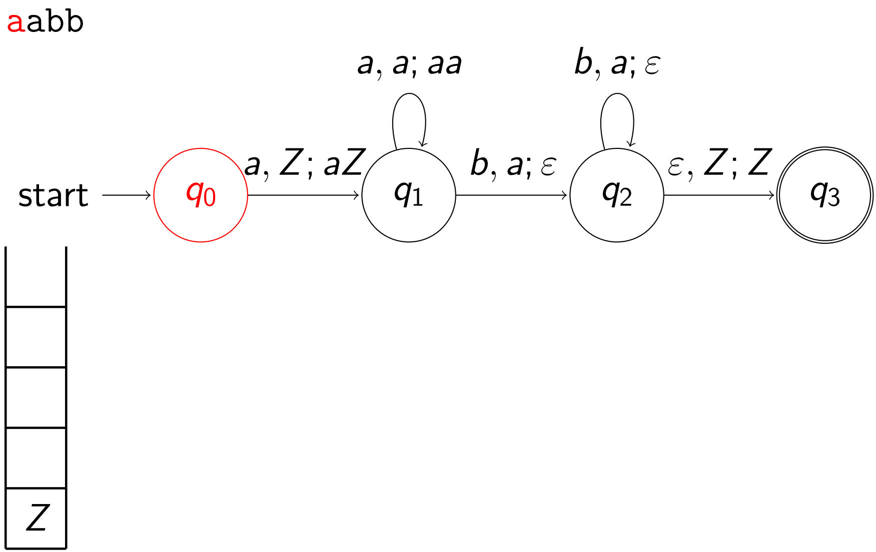
October 3, 2014

PDA for $\{a^n b^n : n > 0\}$

aabb

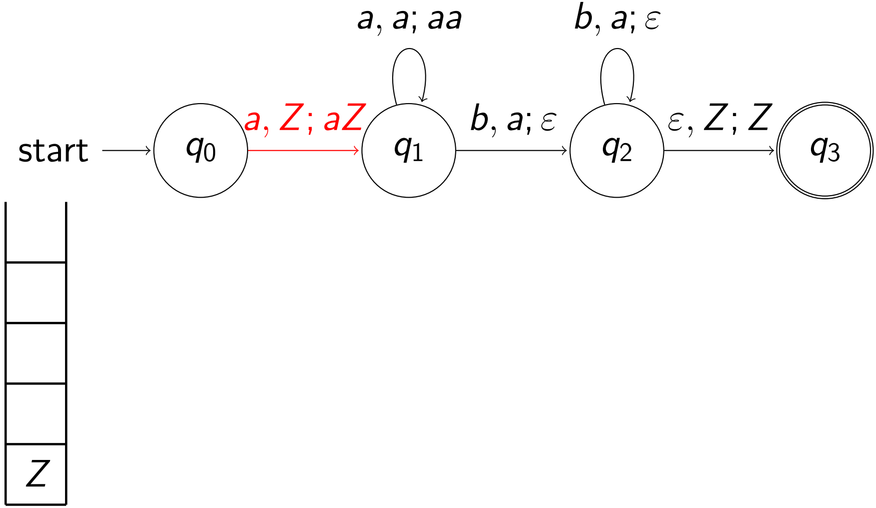


PDA for $\{a^n b^n : n > 0\}$



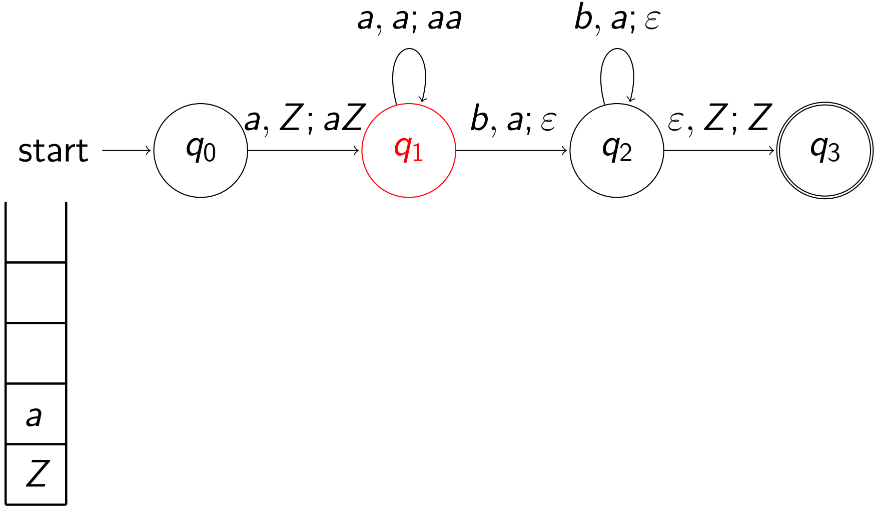
PDA for $\{a^n b^n : n > 0\}$

aabb

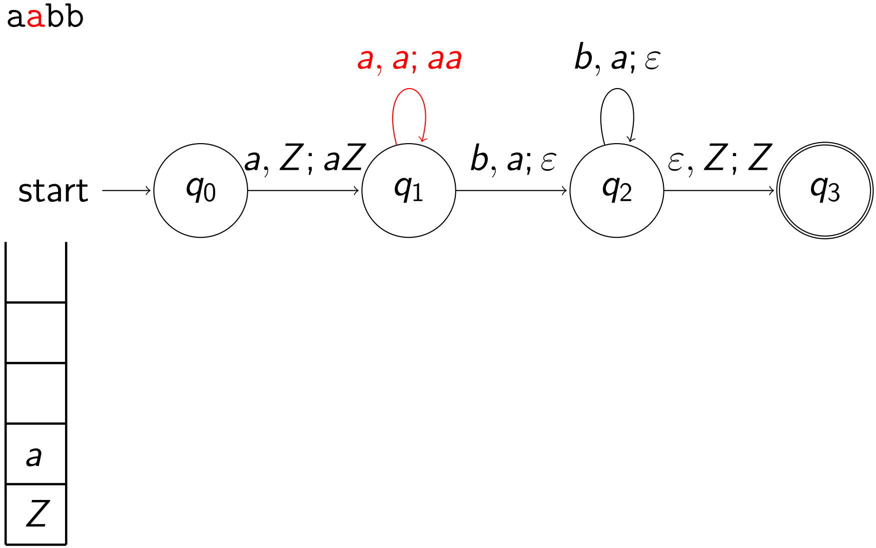


PDA for $\{a^n b^n : n > 0\}$

aabb

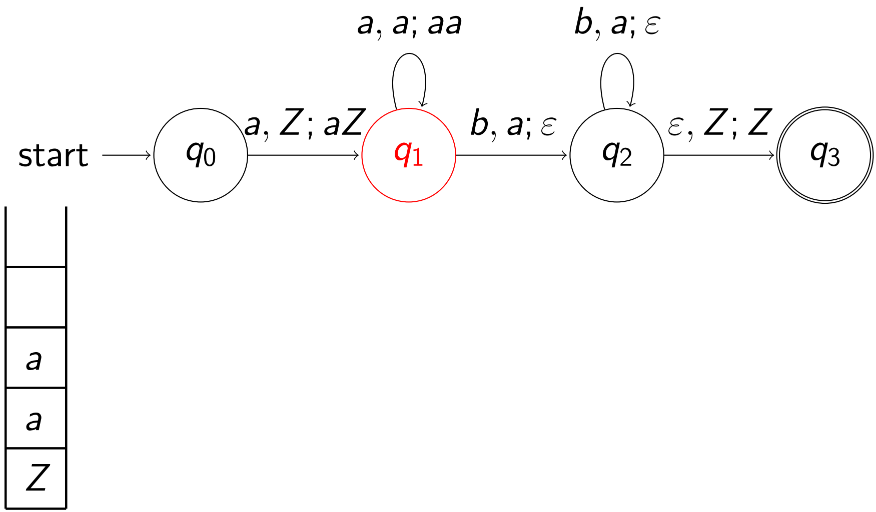


PDA FOR $\{a^n b^n : n > 0\}$

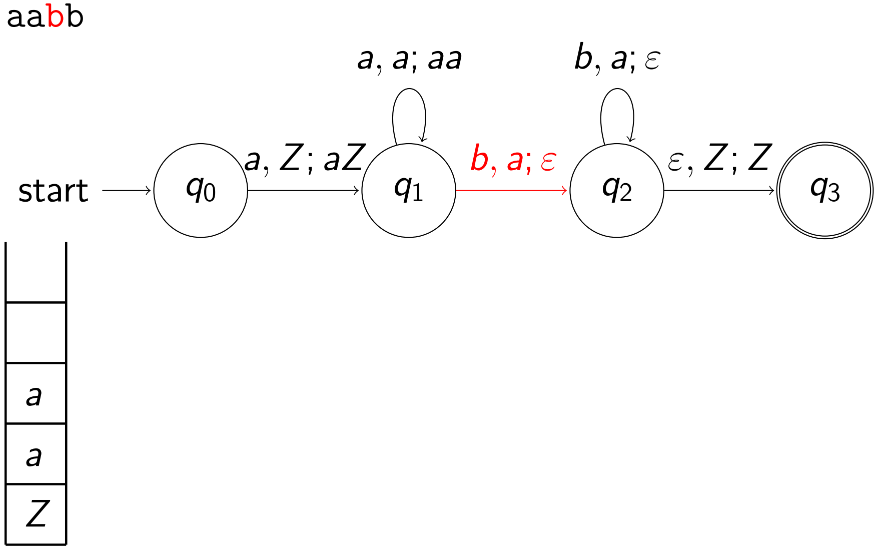


PDA for $\{a^n b^n : n > 0\}$

aabb

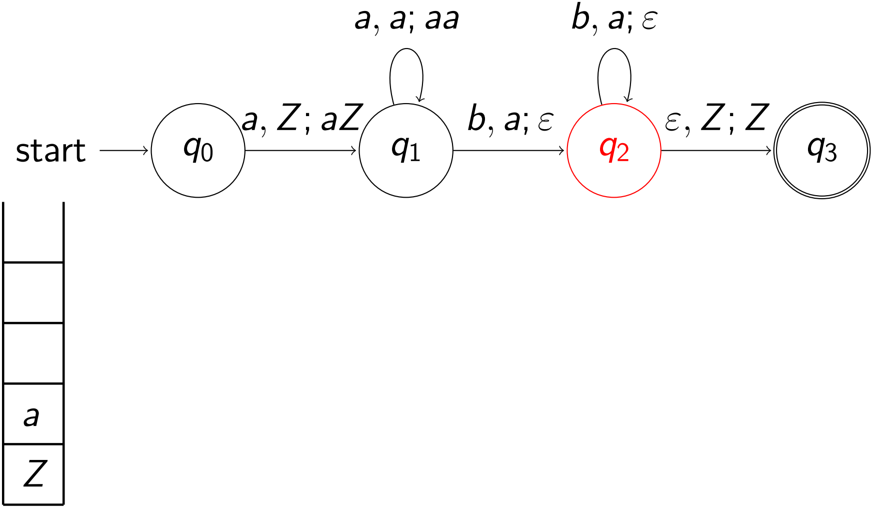


PDA for $\{a^n b^n : n > 0\}$



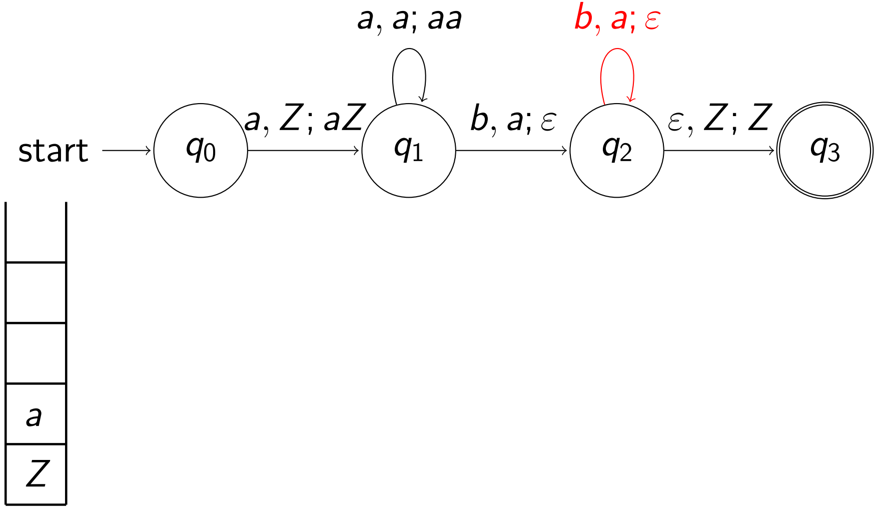
PDA FOR $\{a^n b^n : n > 0\}$

aabb



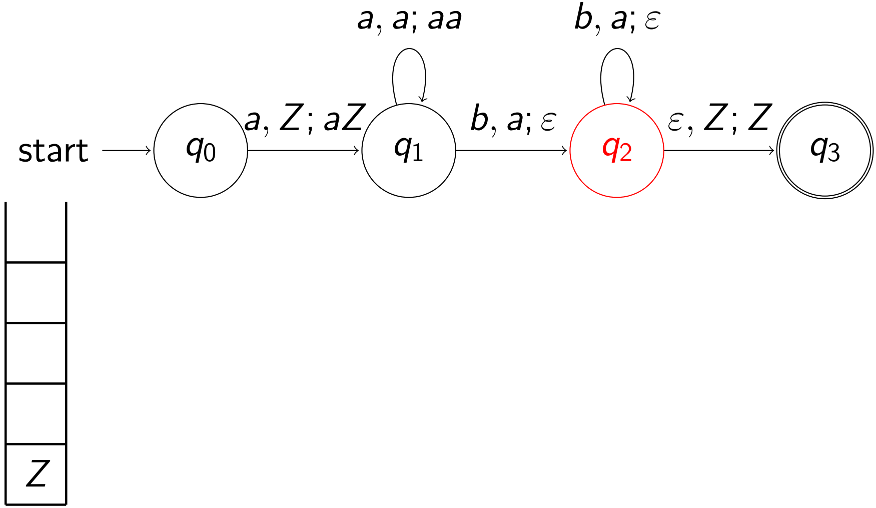
PDA for $\{a^n b^n : n > 0\}$

aabb



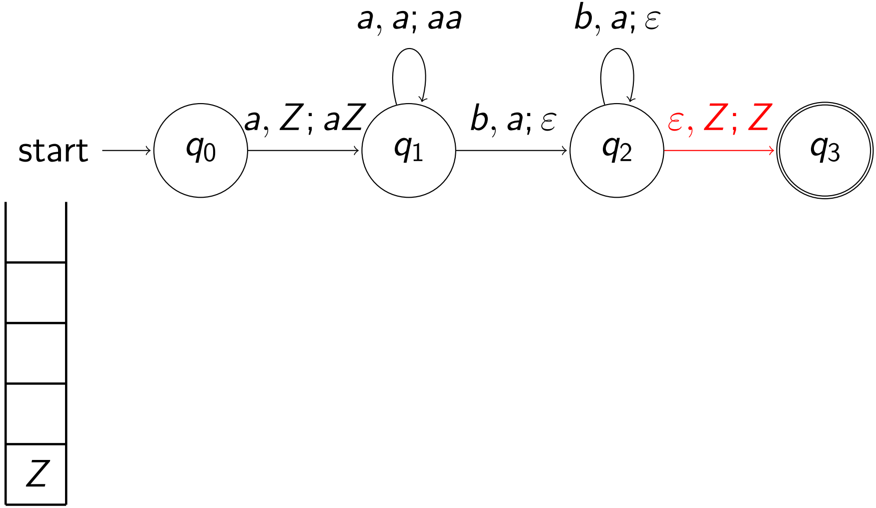
PDA for $\{a^n b^n : n > 0\}$

`aabb`



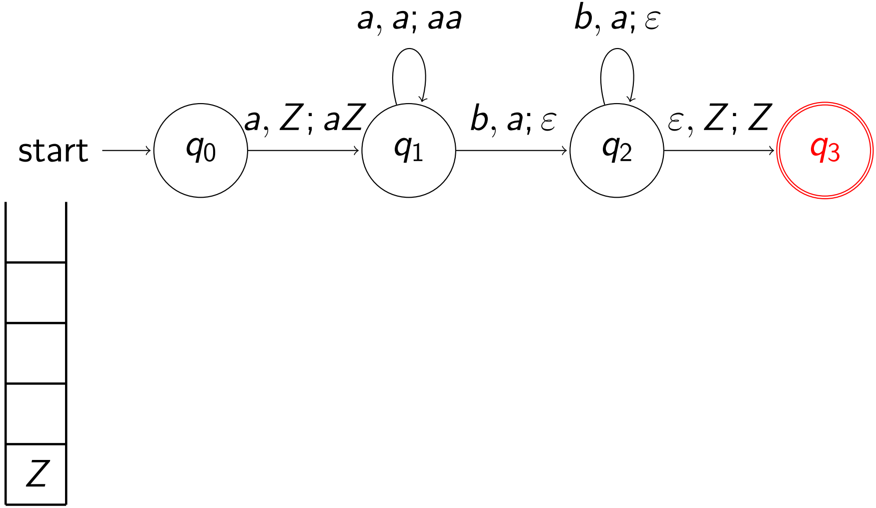
PDA FOR $\{a^n b^n : n > 0\}$

`aabb`



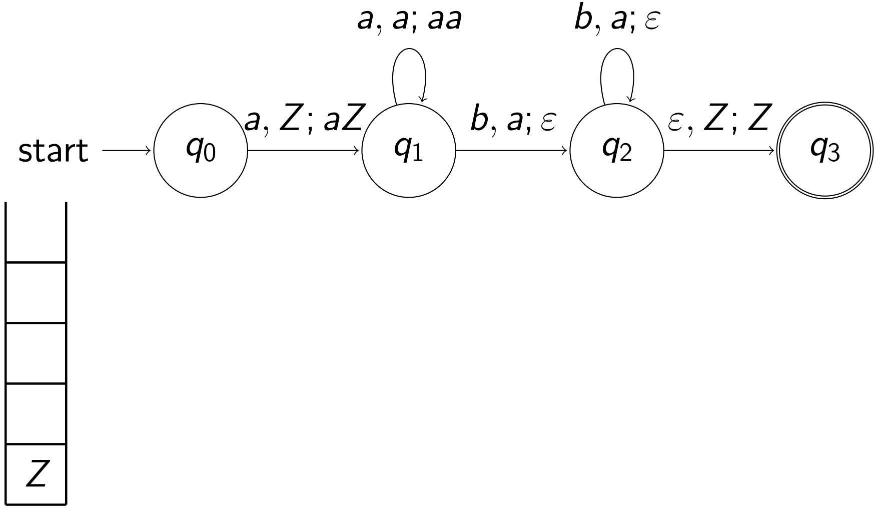
PDA for $\{a^n b^n : n > 0\}$

`aabb`



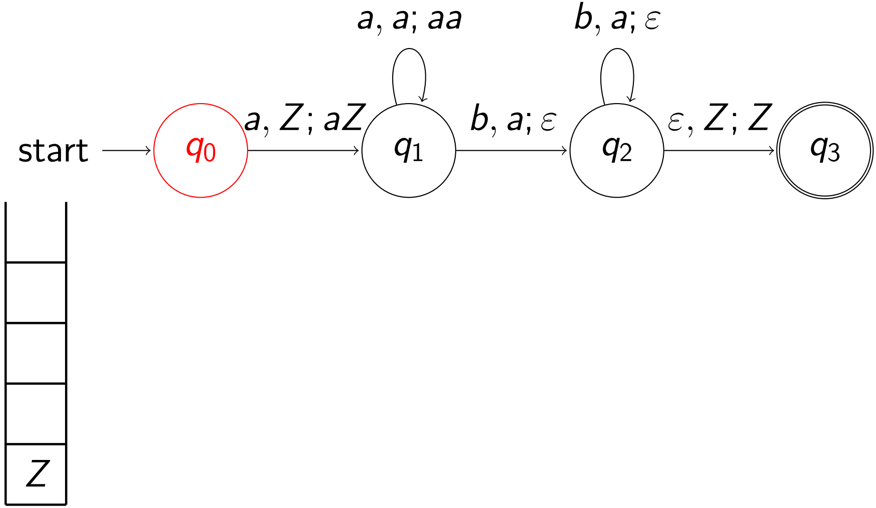
PDA FOR $\{a^n b^n : n > 0\}$

aab



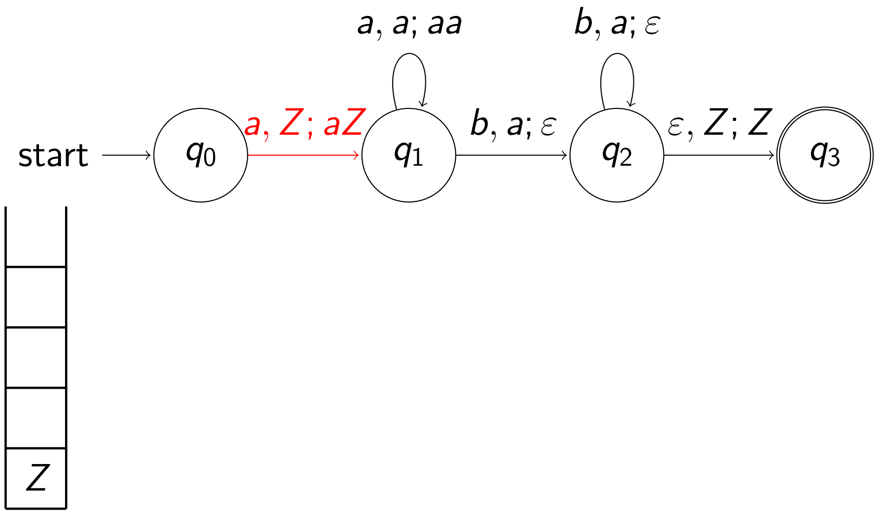
PDA FOR $\{a^n b^n : n > 0\}$

aab



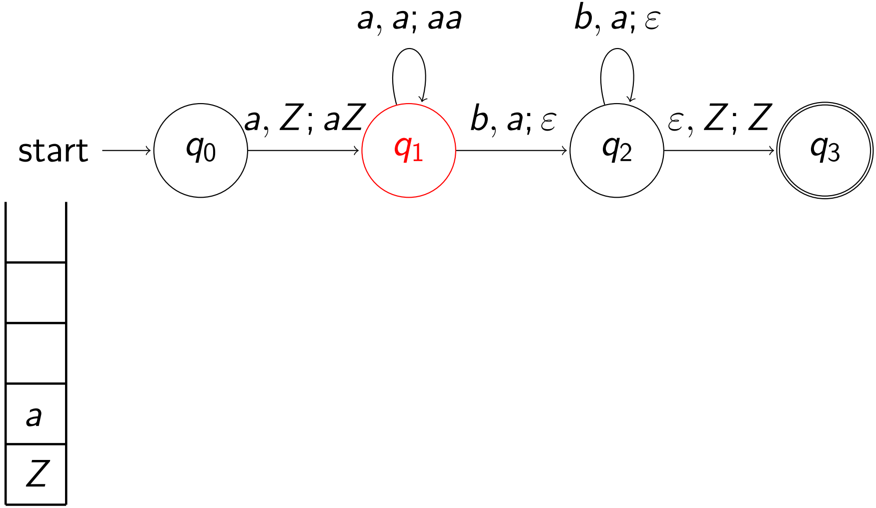
PDA FOR $\{a^n b^n : n > 0\}$

aab

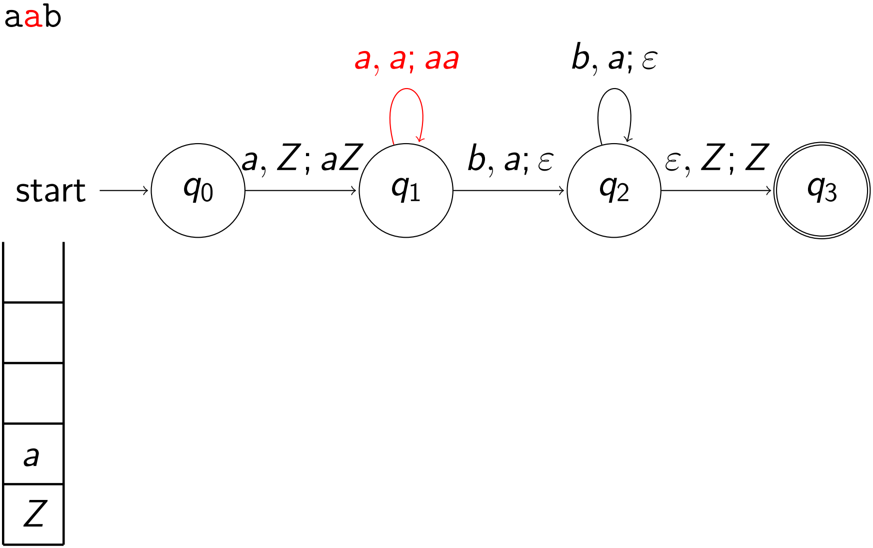


PDA FOR $\{a^n b^n : n > 0\}$

aab

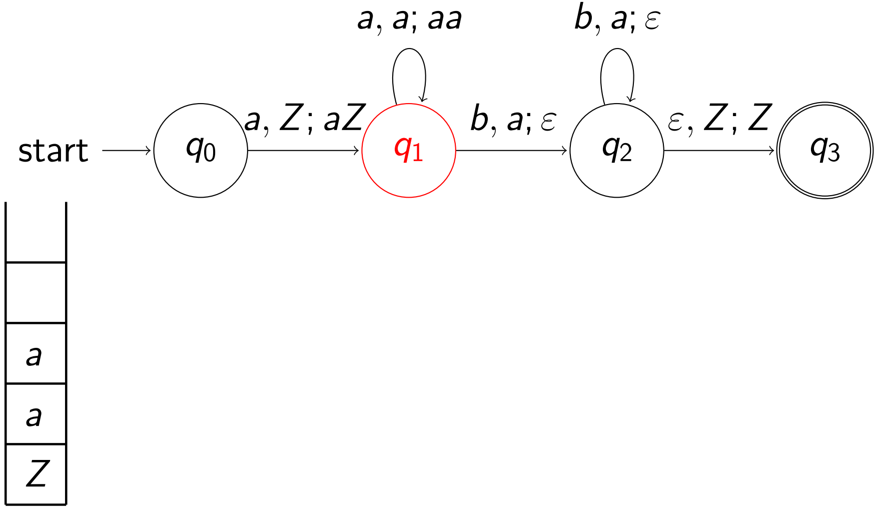


PDA FOR $\{a^n b^n : n > 0\}$



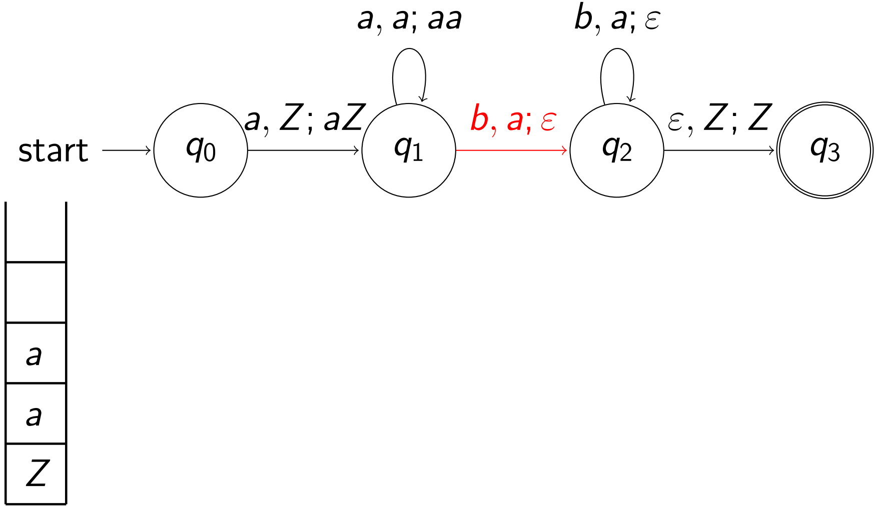
PDA FOR $\{a^n b^n : n > 0\}$

aab



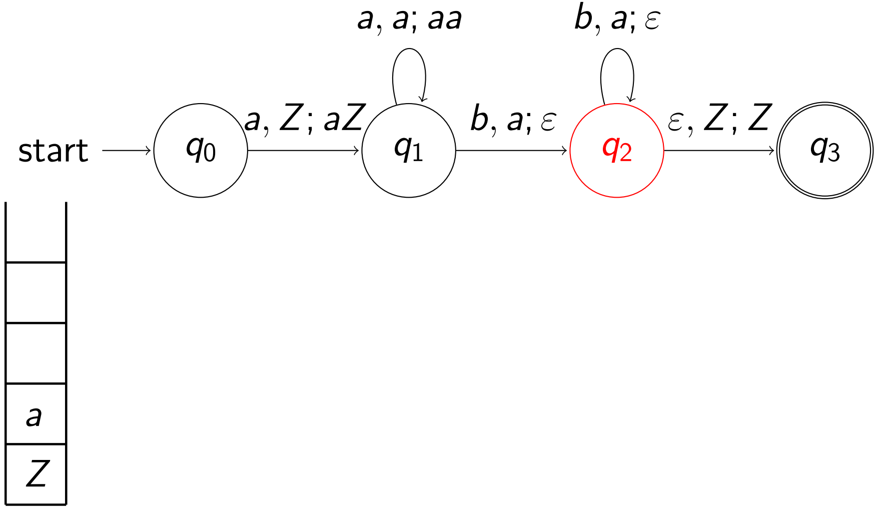
PDA for $\{a^n b^n : n > 0\}$

aab



PDA for $\{a^n b^n : n > 0\}$

aab



Context-Free Grammars (CFG)

- A grammar is a set of *string substitution rules* for producing a set of strings.
- Example: The context-free grammar that generates the language $\{a^n b^n : n > 0\} = \{ab, aabb, aaabbb, ...\}$ is shown below:

$$\begin{array}{lll} S & \rightarrow ab & \text{(base case)} \\ S & \rightarrow aSb & \text{(recursive rule)} \end{array}$$

- The grammar can also be shorten by combining rules with the same left-hand side and using "$|$"
  $S \rightarrow ab \mid aSb$

Derivation

A string x can be derived from a grammar if x can be generated by successive applications of the production rules starting from the start symbol.

Example

Grammar:

$S \rightarrow ab$ (base case)

$S \rightarrow aSb$ (recursive rule)

Derive: `aaabbb`

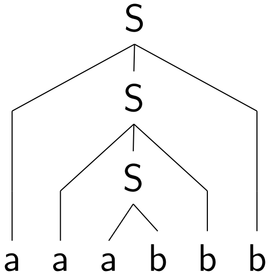
$S \Rightarrow aSb \Rightarrow aaSbb \Rightarrow aaabbb$

Parse Trees (Derivation Tree)

A parse tree is a tree with the *start symbol as the root*, and the *target string forming the leaves* of the tree

Grammar: $S \rightarrow ab \mid aSb$

Derive: `aaabbb`



Context-Free Languages

- All strings that can be generated constitute the *language of the grammar*
- We write $L(G)$ for the language of grammar G
- Any language that can be generated by some context-free grammar is called a *context-free language*

English Language Example

SENTENCE	$\to$	NOUN-PHRASE VERB-PHRASE
NOUN-PHRASE	$\to$	CMPLX-NOUN
	$\to$	CMPLX-NOUN PREP-PHRASE
VERB-PHRASE	$\to$	CMPLX-VERB
	$\to$	CMPLX-VERB PREP-PHRASE
PREP-PHRASE	$\to$	PREP CMPLX-NOUN
CMPLX-NOUN	$\to$	ARTICLE NOUN
CMPLX-VERB	$\to$	VERB $\mid$ VERB NOUN PHRASE
ARTICLE	$\to$	a $\mid$ the
NOUN	$\to$	boy $\mid$ girl $\mid$ flower
VERB	$\to$	touches $\mid$ likes $\mid$ sees
PREP	$\to$	with

English Language Example

Sample strings we can derive from the grammar are:

- `a boy sees`
- `the boy sees a flower`
- `a girl with a flower likes the boy`

Try deriving them using the grammar

English Language Example

Derive: `a boy sees`

$$
\begin{aligned}
\text{SENTENCE} &\Rightarrow \text{NOUN-PHRASE VERB-PHRASE}\\
&\Rightarrow \text{CMPLX-NOUN VERB-PHRASE}\\
&\Rightarrow \text{ARTICLE NOUN VERB-PHRASE}\\
&\Rightarrow \text{a NOUN VERB-PHRASE}\\
&\Rightarrow \text{a boy VERB-PHRASE}\\
&\Rightarrow \text{a boy CMPLX-VERB}\\
&\Rightarrow \text{a boy VERB}\\
&\Rightarrow \text{a boy sees}
\end{aligned}
$$

Formal Definition of CFG

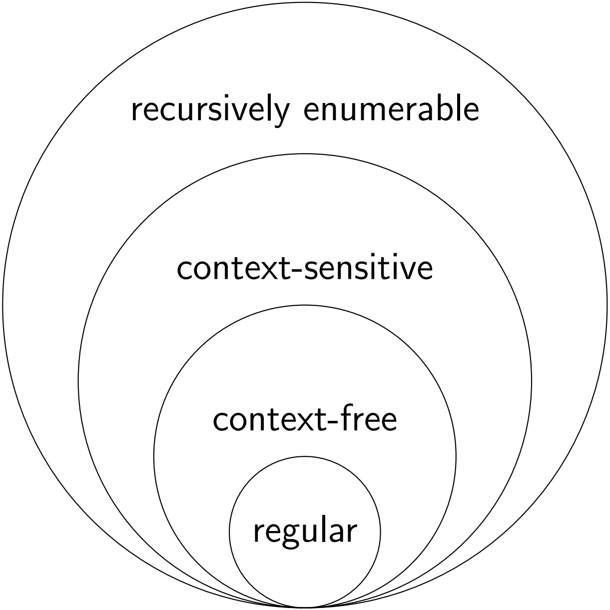
- A *context-free grammar* is a 4-tuple (V, Σ, R, S), where
  - V is a finite set of *variables* (or non-terminals)
  - Σ is a finite set of *terminals*
  - R is a finite set of *rules*
  - $S \in V$ is the start variable.
- The rule for the *rules* is $V \rightarrow (V + T)^*$
- Previous grammar is more formally defined as $G = (\{S\}, \{a, b\}), \{S \rightarrow ab, S \rightarrow aSb\}, S$

Chomsky Hierarchy

by Noam Chomsky
Containment hierarchy of classes of formal grammars

- Regular grammars
  (simplest, weakest)
  - $V \to T^*(V + \varepsilon)$
- Context-free grammars
  - $V \to (V + T)^*$
- Context-sensitive grammars
- Unrestricted grammars/Recursively enumerable grammars
  (most expressive)

Chomsky Hierarchy



References

- Previous slides on CMSC 141
- M. Sipser. Introduction to the Theory of Computation. Thomson, 2007.
- J.E. Hopcroft, R. Motwani and J.D. Ullman. Introduction to Automata Theory, Languages and Computation. 2nd ed, Addison-Wesley, 2001.
- E.A. Albacea. Automata, Formal Languages and Computations, UPLB Foundation, Inc. 2005
- JFLAP, www.jflap.org
- Various online LaTeX and Beamer tutorials