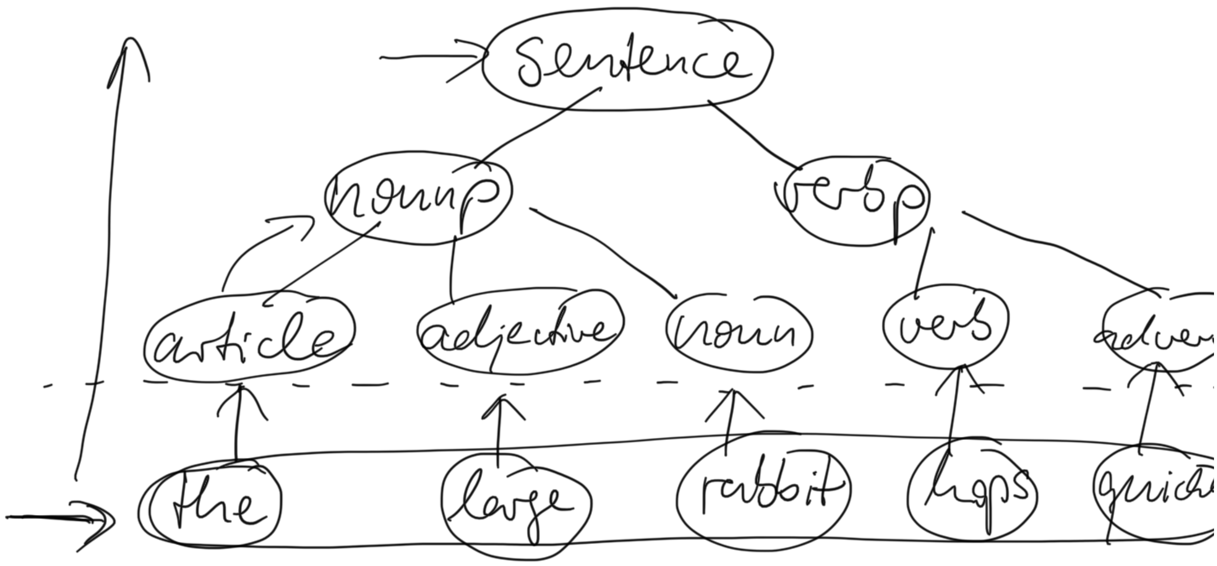
DIT-022 - Lecture 2020-09-07

1) Sentence → nounp verbp
2) nounp → article adjective noun
3) nounp → article noun
4) verbp → verb adverb
5) verbp → verb
6) article → a | the
7) adjective → large | hungry
8) noun → rabbit | mathematician
9) verb → eats | hops
10) adverb → quickly | wildly



Def 1: - alphabet V is a finite non-empty set of elements called symbols

- word over V is a string

word over V is a string of finite length of elements of V

- empty string: λ
- all words: V^*
- grammar: rules for produce words

Def 12 phrase-structure grammar

$$G = (V, T, S, P)$$

V = alphabet

T = set of terminals (subset of V)

$S \in V$

P = set of production rules

$N = V \setminus T$

Def 13: $G = (V, T, S, P)$

$$w_0 = \underline{l}\ z_0\ \underline{r}$$

$$w_1 = \underline{l}\ z_1\ \underline{r}$$

$$z_0 \to z_1 \qquad w_0 \to w_1$$

Example $G = (V, T, S, P)$

$V = \{a, b, A, B, S\}$

$T = \{a, b\}$

$N = V \setminus T = \{A, B, S\}$

$S \in V$

$P = \quad S \to ABa$
$\quad\quad A \to BB$
$\quad\quad B \to ab$
$\quad\quad AB \to b$

$ba \in L(G)$

$BB \notin L(G)$

$S \to ABa \to BBBa$
$\quad \to \underbrace{ab\ ab\ ab\ a}$

ba? $\quad S \to ABa \to ba$ ✓

ab? $\quad S \to \cdots \underline{a}$

<u>Def 14</u>: Let $G = (V, T, S, P)$

$L(G)$ = language of G

= set of <u>all</u> strings of <u>terminals</u> that are derivable from S

$$L(G) = \{ w \in T^* \mid S \stackrel{*}{\Rightarrow} w \}$$

<u>Example</u>: $G = (V, T, S, P)$

$V = \{ S, 0, 1 \}$

$\to T = \{ 0, 1 \}$

$S \quad ; \quad P = \quad S \to 11S$
$\quad\quad\quad\quad\quad\quad S \to 0$

$L(G)$?

$S \to 0.$ $\quad S \to 11S \to 11\,11S$
$\quad\quad \to 11\,11\,0.$

$1s = 111111$

$1's = 1111$

$L(G) = \ldots$

$L(G)$ - set of all strings
that begin with an
<u>even</u> number of 1s

Provide a phrase-structure grammar to generate the set: $\{0^n 1^n \mid n = 0, 1, 2, \ldots\}$

example: $\{\lambda, 01, 0011, 000111, \ldots\}$

$G = (V, T, S, P)$

$V = \{S, 0, 1, \lambda\}$

$T = \{0, 1, \lambda\}$

$P = \quad S \rightarrow \lambda$

$\quad\quad S \rightarrow 0S1$

$S \rightarrow 0\underset{\uparrow 0S1}{S}1 \rightarrow 00\underset{\uparrow \lambda}{S}11 \rightarrow 0011.$

<u>Example</u>: $\{0^m 1^n \mid m, n \in \mathbb{N}_0\}$ $= \{0, 1, 2, 3, \ldots\}$

$\Longrightarrow \{\lambda, 1, 111, 1111, 00, 0, 01, \ldots\}$

$G = (V, T, S, P)$

$V = \{S, 0, 1, \lambda, A, B\}$

$T = \{0, 1, \lambda\}$

S

P: $S \rightarrow AB$

$A \rightarrow 0A$

$A \rightarrow \lambda$

$B \rightarrow 1B$

$S \rightarrow AB$
$\rightarrow 0AB$
$\rightarrow 00AB$
$\rightarrow 00B$
$\rightarrow 001B$

$B \to 1B \qquad \to 001B$

$B \to \lambda \qquad \to 001.$

Show that "cbab" belongs to the language generated by

$G = (V, T, S, P)$

$V = \{a, b, c, A, B, C, S\}$

$T = \{a, b, c\}$

$S \quad P = \quad S \to AB$

$A \to Ca$

$B \to Ba$

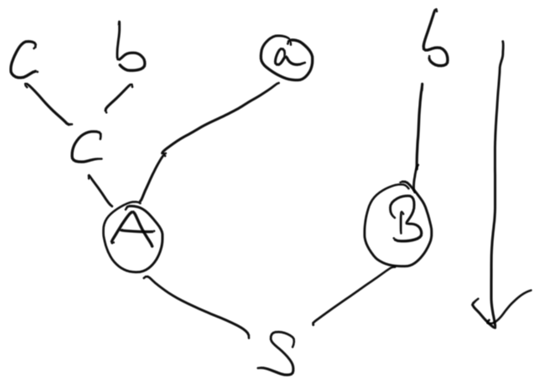
$B \to Cb$

$B \to b$

$C \to cb$

$C \to b$

$cbab \in L(G)$



parsers

CYK

LL(k) parser

EBNF

$S \to AB$	$S \to A\ B$
	S = A, B ;
$A \to Ca$	A = C, "a" ;
$B \to Ba$	B = B, "a" | C, "b" | "b" ;
$B \to Cb$	
$B \to b$	

definition =

definition =

concatenation ,

termination ;

alternation |

optional [...]

repetition { ... }

terminal " "

$*$, $+$

$A \to c^*$ ($0 \ldots n$): $\lambda, c, ccc, \ldots$

$A \to c^+$ ($1 \ldots n$)

---

finite state machines

Def 15 FSM

$$M = (S, I, O, f, g, s_0)$$

S = states

I = input alphabet

O = output alphabet

f = transition function

g = assigns output to input

s_0 = starting state

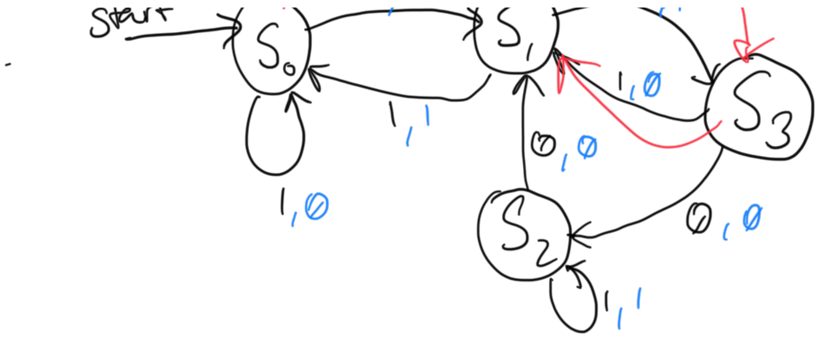
Example $S = \{s_0, s_1, s_2, s_3\}$

$I = \{0, 1\}$ $O = \{0, 1\}$

State table

state	f Input 0	f Input 1	g Input 0	g Input 1
$\to s_0$	s_1	s_0	1	0
$\to s_1$	s_3	s_0	1	1
$\to s_2$	s_1	s_2	0	1
$\to s_3$	s_2	s_1	0	0

transition output



$\emptyset\,\emptyset\,1 \rightarrow$

$1\,1\,\emptyset$

$$X = x_1\, x_2\, x_3 \cdots x_k$$

S_0

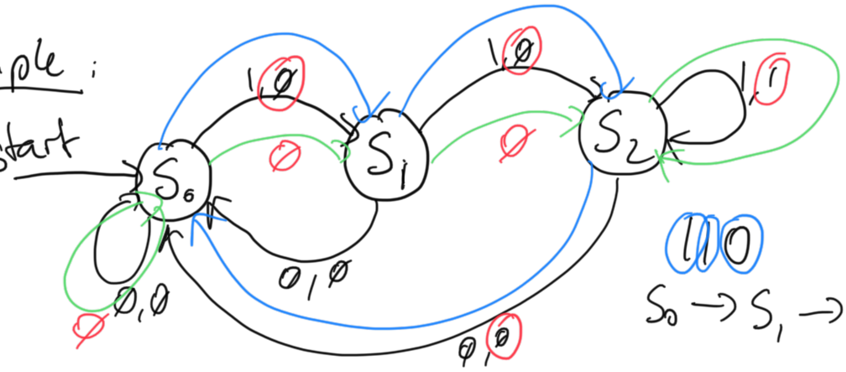
transition
$$\begin{cases} s_1 = f(s_0, x_1) \\ s_2 = f(s_1, x_2) \\ \quad \vdots \\ s_j = f(s_{j-1}, x_j) \quad j = 1,2,\dots,k \end{cases}$$

output
$$\begin{cases} y_1 = g(s_0, x_1) \\ y_2 = g(s_1, x_2) \\ \quad \vdots \\ y_j = g(s_{j-1}, x_j) \quad j = 1,2,\dots,k \end{cases}$$

Def 16: $M = (S, I, O, f, g, s_0)$ and $L \subseteq I^*$

M accepts L if an input $x \in L$ produces as last output a 1.

Example:



110

$S_0 \rightarrow S_1 \rightarrow S_2 \rightarrow S_0$

$\emptyset\emptyset\emptyset$

$\emptyset 1 1$

0111
0001

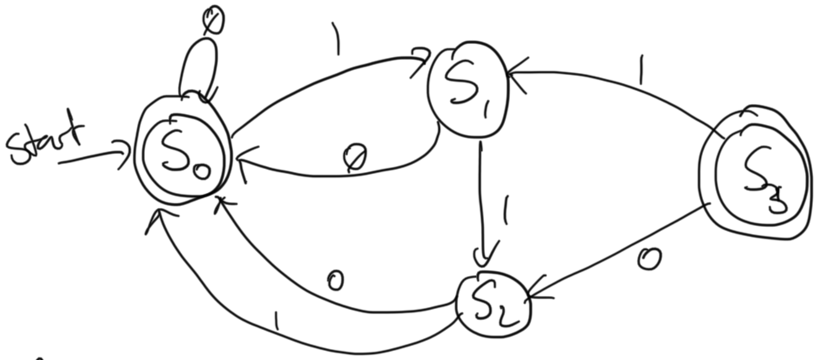
Mealy machines output is produced on the transition

Moore machines output is produced in the state

FSM

Def 17: $M = (S, I, f, s_0, F)$

accepting states



λ 0 10 110 111

11

Def 18: a string x is said to be recognized/accepted by M if it takes the initial state s_0 to a final; i.e. $f(s_0, x) \in F$.

The language accepted by M is denoted by $L(M)$.

Example: provide an FSM

that only accepts
consecutive 1s

λ ✓