INF210 Overview

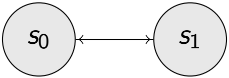
Åsmund Kløvstad

Universitetet i Bergen

December 4, 2021

State Transition Systems

- a tuple (S, R)
- $S = \{s_0, s_1 ...\}$
- $R \subseteq S \times S$, often written $s \rightarrow s'$
- example: switch $= (\{s_0, s_1\}, \{s_0 \times s_1, s_1 \times s_0\})$



Important Properties of STSs

- if a state has no rules going *from* it, it is a **terminal state**.
- if there is a sequence of steps from state a to state b, b is **reachable** from a. Write $R^*(a, b)$.
- if R is a function (in the set theory sense), then STS is **deterministic**.
- we can make a non-deterministic STS deterministic with $(\mathcal{P}(S), R_{\mathcal{P}})$
- if all paths eventually converge the STS is **confluent**.
- switch is deterministic and confluent, and has no terminal states.

Labeled Transition Systems

Syntax

- extends STSs with labels
- a 3-tuple (S, L, R).
- $R \subseteq S \times L \times S$, often written $s \xrightarrow{a} s'$

Reachability

- consider the STS (S_L, R_L) with:
  - $S_L = S \times L^*$
  - $s \times a \cdot \mathbf{w} \rightarrow s' \times \mathbf{w} \in R_L$ iff $s \xrightarrow{a} s' \in R$
- s' is reachable from s by $\mathbf{w}$ if $s' \times \lambda$ is reachable from $s \times \mathbf{w}$.
- write $s \times \mathbf{w} \vdash_R^* s' \times \lambda$

Languages

- language over alphabet: $L \subseteq \Sigma^*$
- ex. $\{a^n b^n \mid n \in \mathbb{N}\}$ over $\Sigma = \{a, b\}$
- deciding membership in a language
- classifying classes of languages

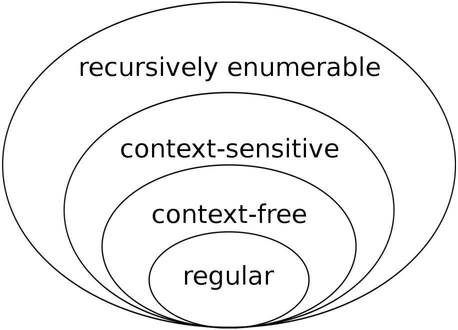


Figure: from `https://en.wikipedia.org/wiki/Chomsky_hierarchy`

Grammars

- $G = (\Sigma, N, S, \mathcal{R})$
- rules are $\mathbf{l} \Rightarrow \mathbf{r}$ with at least one non-terminal in $\mathbf{l}$
- **generates** a language over Σ
- $\mathbf{uav} \Rightarrow_G \mathbf{ubv}$ if $\mathbf{a} \Rightarrow \mathbf{b} \in \mathcal{R}$
- write $\mathbf{u} \Rightarrow_G^* \mathbf{v}$ for $\mathbf{u}, \mathbf{v} \in (\Sigma \cup N)^*$ for "**u** generates **v**"
- $L_G = \{\mathbf{w} \in \Sigma^* \mid S \Rightarrow_G^* \mathbf{w}\}$
- example: $G = (\{a\}, \{S\}, S, \{S \Rightarrow \lambda, S \Rightarrow Sa\})$ generates a^*

Machines

- FA/FSM's, PDA's, TM's
- **accepts** input words
- $L_M = \{$words accepted by M$\}$

Finite Automata

- $M = (\Sigma, Q, q_0, \Upsilon, F)$
- a finite LTS with initial and final states
- the language accepted by M is
  $L_M = \{\mathbf{w} \in \Sigma^* \mid q_0 \times \mathbf{w} \vdash^*_{\Upsilon} q_i \times \lambda \text{ and } q_i \in F\}$
- deterministic/non-deterministic are equally powerful
- example accepting a^+b^+

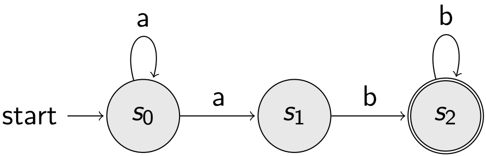


Figure: $M = (\{a, b\}, \{s_0, s_1, s_2\}, s_0, \{(s_0, a, s_0), (s_0, a, s_1), (s_1, b, s_2), (s_2, b, s_2)\}, \{s_2\})$

Regular Languages

- **regular** languages are inductively defined by:
  - $\emptyset, \{\lambda\}$ and $\{a\}$ are regular for $a \in \Sigma$
  - if L, L' regular, then $L \cup L', L \cdot L'$ and L^* are regular
- $\overline{L}$ and $L \cap L'$ are also regular

Kleene's Theorem

Theorem

a language is regular $\iff$ *it is accepted by a finite automata*

$\Rightarrow$

Construct FAs for the empty language, $\{\lambda\}$ and singleton languages. Then construct FAs for union, concatenation, and Kleene star.

$\Leftarrow$

Define $R(i, k, j) =$
$\{\mathbf{w} \mid q_j \text{ is reachable from } q_i \text{ by } \mathbf{w} \text{ without visiting } q_m \text{ with } m \geq k\}$
Then show
$R(i, k+1, j) = R(i, k, j) \cup R(i, k, k) \cdot R(k, k, k)^* \cdot R(k, k, j)$

Pumping Lemma

- if L is regular then there exists some $n > 0$ s.t for any $\mathbf{w} \in L$ longer than n, $\mathbf{w} = \mathbf{xyz}$ with $|\mathbf{y}| \geq 1$ and $|\mathbf{xy} \leq n|$ Then $\mathbf{xy^kz} \in L$ for all $k \geq 0$
- makes sense because some state must be visited twice
- very useful for showing a language is **not** regular

Syntactic Monoid of a Language

- given a language L over Σ
- $LR(\mathbf{w}) = \{\mathbf{x} \times \mathbf{y} \mid \mathbf{xwy} \in L\}$
- $\mathbf{u} \approx \mathbf{v} := LR(\mathbf{u}) = LR(\mathbf{v})$
- this is left-right invariant, so $[\mathbf{u}]_\approx \cdot [\mathbf{v}]_\approx := [\mathbf{uv}]_\approx$ is well defined
- $(\Sigma^*_{/\approx}, \cdot, [\lambda]_\approx)$ is the **syntactic monoid** of L.

Transformation Monoid of a FA

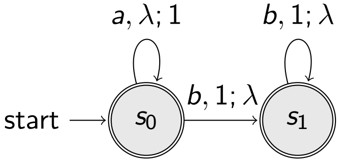
- given $M = (\Sigma, Q, q_0, \Upsilon, F)$
- each $a \in \Sigma$ induces a function $\overline{a} : Q \to Q$ by $q \mapsto \Upsilon(q, a)$
- with $\overline{\lambda} = id_Q$ and $\overline{\mathbf{w}a} = \overline{\mathbf{w}} \circ \overline{a}$ this extends to words
- $\overline{\mathbf{vw}} = \overline{\mathbf{v}} \circ \overline{\mathbf{w}}$, so $\overline{} : \Sigma^* \to (Q \to Q)$ is a (monoid) homomorphism
- call the image $\overline{\Sigma^*} \subseteq (Q \to Q)$ the **transformation monoid** of M
- the syntactic monoid of $L \cong$ the transformation monoid of M_L (intrinsic/minimal FA)

Regular Grammars

- rules are either $A \Rightarrow B$, $A \Rightarrow aB$ or $A \Rightarrow \lambda$ for $A, B \in N, a \in \Sigma$
- generate regular languages
- let FA $M = (\Sigma, N \cup \{f\}, S, \Upsilon, \{A \in N \mid A \Rightarrow \lambda \in \mathcal{R}\})$ with Υ:
  - $A \xrightarrow{a} B$ for $A \Rightarrow aB \in \mathcal{R}$
  - $A \xrightarrow{a} f$ for $A \Rightarrow a \in \mathcal{R}$
- then $L_M = L_G$

Pushdown Automata

- finite automaton with a stack:
  - Σ: an alphabet
  - Γ: a stack alphabet
  - Q: a finite set of states
  - $q_0 \in Q$: an initial state
  - $\Delta \subseteq Q \times (\Sigma_\lambda \times \Gamma_\lambda \times \Gamma_\lambda) \times Q$: a transition relation
  - $F \subseteq Q$: a set of final states
- accepts in final state with empty stack
- example:



Context Free Grammars

- left hand side is a single non-terminal
- generate a superset of regular languages
- call this class **context free** languages
- ex. $S \Rightarrow \lambda, S \Rightarrow aSb$ generates $a^n b^n$

Context Free Languages

Theorem

a language is context free $\iff$ it is accepted by a pushdown automata

$\Leftarrow$

Given PDA $M = (\Sigma, \Gamma, \{s_0\}, s_0, !, \Delta)$ construct $G' = (\Sigma, \Gamma, !, \mathcal{R})$ with $\mathcal{R}$:

- $A \Rightarrow a\mathbf{W}$ for each $(a, A; \mathbf{W}) \in \Delta$
- $A \Rightarrow \mathbf{W}$ for each $(\lambda, A; \mathbf{W}) \in \Delta$

note: empty stack acceptance, start with ! on stack, allow pushing words

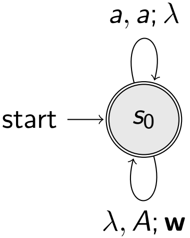
Context Free Languages

Theorem

a language is context free $\iff$ *it is accepted by a pushdown automata*

$\Rightarrow$

Given CFG $G = (\Sigma, N, S, \mathcal{R})$ construct:



for each $a \in \Sigma$ and $A \Rightarrow \mathbf{w} \in \mathcal{R}$

Parsing

- given a $\mathbf{w} \in \Sigma^*$, identify syntactic structure
- top-down: attempt to generate $\mathbf{w}$ from S (PDA)
- bottom-up: split into subwords and find rules to generate them

$$
\begin{array}{ccccccccc}
 & & & & \{S\} & & & & \\
 & & & \emptyset & & \{A\} & & & \\
 & & \{S\} & & \emptyset & & \{S\} & & \\
 & \emptyset & & \{A\} & & \emptyset & & \{B\} & \\
\{S\} & & \{Z_+\} & & \{S\} & & \{Z_*\} & & \{S\} \\
1 & & + & & 2 & & \star & & 3
\end{array}
$$

Figure: from `Parsing.pdf`

Pumping Lemma for CFLs

- given CFG $G = (\Sigma, N, S, \mathcal{R})$
- any $\mathbf{s} \in L_G$ longer than $2^{|N|-1}$ can be written as $\mathbf{s} = \mathbf{uvwxy}$ with $|\mathbf{vx}| \geq 1$ s.t $\mathbf{uv}^n\mathbf{wx}^n\mathbf{y} \in L_G$ for all $n \in \mathbb{N}$

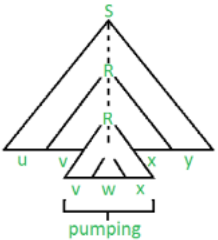


Figure: from https://www.geeksforgeeks.org/pumping-lemma-in-theory-of-computation/

Turing Machine

- $M = (Q, \Sigma, q_0, \delta, H)$ with:
  - Q: a finite set of states
  - Σ: a finite alphabet which includes the blank symbol $\#$
  - $q_0 \in Q$: a starting state
  - $\delta : (Q \setminus H) \times \Sigma \to Q \times \Sigma \times \{-1, 0, 1\}$: a (possibly partial) transition function
  - $H = \{h_0, h_1\}$: a rejecting and an accepting halting state
- a head with state reading an infinite tape
- accepts a word **w** iff it halts in h_1 when started with **w** on the tape
- **decides** a language only if h_0 is reached for $\mathbf{w} \notin L$

Phrase-Structure Grammars

- grammars with rules $\mathbf{u} \Rightarrow \mathbf{w}$ s.t $\mathbf{u}$ has at least one non-terminal
- as powerful as Turing Machines

Create a PSG which work like:

1. generate $\{\mathbf{w}!q_0H\mathbf{w}!\}$
2. maintain $\mathbf{w}$! [tape] [state] H [tape] !
3. when state is h_1, remove everything between !'s (and then the !'s)

Halting Problem

Theorem

Standardize and enumerate all Turing Machines $\{T_i\}$. *Let*
$L = \{1^i01^j \mid T_i \text{ halts on input } 1^j\}$
Then there is no Turing Machine which decides L

- assume M decides L
- construct M' s.t it acts on input 1^i:
  1. change tape to 1^i01^i
  2. move to first cell
  3. emulate M on the tape
  4. if M accepts, loop infinitely. if M rejects, accept
- enumerate $M' = T_k$ and run M' with 1^k as input
- if M accepts, then M' loops forever on 1^k, so M must reject
- if M rejects, then M' halts on 1^k, so M must accept

Splicing Languages

- abstracts DNA as a word over $\Sigma = \{a, c, g, t\}$
- splicing rules: $(\mathbf{u}, \mathbf{u}', \mathbf{v}, \mathbf{v}') \in (\Sigma^*)^4$
- splicing action: split between **u,u'** and **v,v'**. Recombine **uv'**
- splicing system: (Σ, R, I) with an alphabet, set of rules, finite initial language
- language generated by system: smallest closed language
- L is a **splicing language** if it is generated by some splicing system
- given a regular language and a set of splicing rules, we can check if the language respects the rules