## Chapter 1

Finite Automata ($Q-$states, $\Sigma-$alphabet, $\delta-$transitions, $q_0-$start, $F \subset Q-$accept). Language is **regular** if a finite automaton recognizes it. Two machines are equivalent if they recognize the same language.

- Derterministic (DFA) Restrict to one transition for each unique symbol.
- Nondeterministic (NFA) Every NFA has an equivalent DFA and any DFA is a valid NFA. Therefore a language is **regular** if and only if some NFA recognizes it.
- DFA to NFA Start at start state(s). Follow and write next possible states per symbol. Create new row for resulting states. Repeat until no new states. Should be 1 more state than the NFA.
- Generalized nondeterministic finite automaton (GNFA) Only one start and reject state. Transitions are regular expressions. Used to convert DFA to a RE.
- DFA to DE Add new start state and accept state. Transition start to old start and from old accept to new accept. Identify destination states from the state that will be removed. Identify all paths destination states have that go through the state that will be removed. Write new transitions excluding the removed state. Repeat.

Regular Languages are closed under **union, intersection, complement, concatenation, star (*)**. All finite languages are Regular Languages.

Power Set is the set of all subsets of a language. Size of $P(A) = 2^{|A|}$.

Regular Expression. R is a RE if it is (1) a character in the alphabet associated with R. (2) the empty string. (3) the empty language. (4) two regular languages under union. (5) two regular languages under concatenation. (6) a regular language under star. **Order of Operations** is parenthesis, star, concatenation, union. A language described by a RE is **regular**.

Pumping Lemma for RL A string of length at least pumping-length can be broken up into $xyz$ such that (1) $xy^iz$ is in the language for any $i \geq 0$. (2) $|y| > 0$ (3) $|xy| \leq p$.

Finite Automata Theorems For a finite automata $M$ with $n$ states (1) $L(M)$ is non-empty if and only if $M$ accepts a string of length less than $n$ (2) $L(M)$ is infinite if and only if $M$ accepts a string of length $i$ where $n \leq i < 2n$. It is possible to create a FA that can determine if two FA are equivalent and taking a finite amount of time if they are equivalent.

## Chapter 2

Context-free Grammar ($V-$variables (states), $\Sigma-$terminals (symbols), $R-$rules (transitions), $S-$start). Parse-trees show the path the CFG takes to output the string. Any language made by a CFG is a **context-free** language. A CFG is **ambiguous** if there is more than one way to generate a string (two parse trees). A CFL is **inherently ambiguous** if all grammars for the language are ambiguous. **Leftmost** deviation means the leftmost remaining variable is the one replaced; same for rightmost.

Context-free Languages are closed under **union, concatenation, star (*)**. All Regular languages are context-free.

Chomsky Normal Form if every rule is of the form $A \to BC$ or $A \to a$. The start variable can have a $\epsilon$. (1) Add new start variable with rule to old start variable. (2) Eliminate all $\epsilon$ rules. (3) Eliminate all unit rules. (3) Convert remaining to proper form by moving stuff around. Any CFL can be generated by a CFG in Chomsky normal form.

Pushdown Automata (PDA) Same setup as a FA, except the inclusion of a stack and the transitions that can pop or push something on the stack. A language is **context-free** if and only if some PDA recognizes it. Every regular language is context-free.

Pumping Lemma for CFL If L is a CFL, then there is a pumping-length where if a string in L is at least pumping-length then the string can be broken up into $uvxyz$ where (1) $uv^ixy^iz \in L$ for all $i \geq 0$. (2) $|xy| > 0$ (3) $|vxy| \geq p$

## Examples

Prove $\{0^n1^m0^n \mid m,n \geq 0\}$ is not regular. Assume this language ($A$) is regular, so then there must exist a variable $p$, the pumping length. Choose $w = 0^p10^p$ as the test word. $|w| > p$ and $w \in A$. As $|xy| \leq p$, $x$ and $y$ must be composed of only zeros. Additionally, as $|y| > 0$, $y$ would then have to equal $0^k$ for some $k > 0$. For $xy^iz$, choose $i = 0$ and the resulting word should still be in $A$. However $xy^0z = xz = 0^{p-k}10^p$. This resulting word is not in $A$ therefore our assumption was incorrect.

Prove $\{a^nb^m \mid m \leq n^2\}$ is not context-free. Choose $S = a^{p+1}b^{p^2+1}$ There are then three cases for $vxy$. (1) $a^{p+1}$, (2) $a^pb^{p^2}$, (3) $b^{p^2+1}$. We pump down on case 1 and 2, and pump up for case 3. For case 1, the number of b's is greater than the number of a's squared ($a^{p+1-1}b^{p^2+1}$). For case 2, the number of a's and b's become equal ($ab$), which is not what the language wants. For case 3, the number of a's squared will be greater than the number of b's. ($a^{p+1}b^{p^2+1+1}$). With all options exhausted, the language cannot be context-free.

Convert the following CFG to Chomsky normal form. (S is already new start state)

| | | |
|---|---|---|
| $S \to aAA \mid aBC \mid abc$ | $S \to DAA \mid DBC \mid DEF$ | $S \to DI \mid DH \mid DJ$ |
| $A \to AA \mid Aa \mid ab$ | $A \to AA \mid AD \mid DE$ | $A \to AA \mid AD \mid DE$ |
| $B \to aaBC \mid BC$ | $B \to DDBC \mid BC$ | $B \to GH \mid BC$ |
| $C \to a \mid bc$ | $C \to a \mid EF$ | $C \to a \mid EF$ |
| | $D \to a$ | $D \to a$ |
| | $E \to b$ | $E \to b$ |
| | $F \to c$ | $F \to c$ |
| | | $G \to DD$ |
| | | $H \to BC$ |
| | | $I \to AA$ |
| | | $J \to EF$ |

Removing $\epsilon$ rules from CFGs

| | | | | |
|---|---|---|---|---|
| $A \to B \mid C$ | $S \to A$ | $S \to A \mid \epsilon$ | $S \to A \mid \epsilon$ | $S \to A \mid \epsilon$ |
| $B \to aCa \mid \epsilon$ | $A \to B \mid C$ | $A \to B \mid C$ | $A \to B \mid C$ | $A \to B \mid C$ |
| $C \to bAb \mid \epsilon$ | $B \to aCa \mid \epsilon$ | $B \to aCa \mid aa$ | $B \to DCD \mid DD$ | $B \to DF \mid DD$ |
| | $C \to bAb \mid \epsilon$ | $C \to bAb \mid bb$ | $C \to EAE \mid EE$ | $C \to GE \mid EE$ |
| | | | $D \to a$ | $D \to a$ |
| | | | $E \to b$ | $E \to b$ |
| | | | | $F \to CD$ |
| | | | | $G \to EA$ |