MA/CSSE 474 Day 18 Summary

Shorthand notation: $S \to \varepsilon \mid aT \mid bT$ $T \to a \mid b \mid aS \mid bS$

1. CFG (context-free grammar) formal definition
   a. CFG G = (V, Σ, R, S), (each part is finite)
      i. **Σ** is the ***terminal alphabet***; it contains the symbols that make up the strings in $L(G)$, and
      ii. *N is the **nonterminal alphabet*** a set of working symbols that G uses to structure the language. These symbols disappear by the time the grammar finishes its job and generates a string. **(Note:** $\Sigma \cap N = \varnothing$.)
      iii. Rule alphabet: **V** = $\Sigma \cup N$
      iv. **R**: A set of rules (a.k.a. productions) of the form $A \to \beta$, where $A \in N$ and $\beta \in V^*$.
      v. G has a unique **start symbol**, **S** $\in$ N

2. Formal definition of *derivation* and related things:
   a. $x \Rightarrow_G y$ iff $x = \alpha A \beta$, $y = \alpha \gamma \beta$, and $A \to \gamma$ is in R
   b. $w_0 \Rightarrow_G w_1 \Rightarrow_G w_2 \Rightarrow_G \ldots \Rightarrow_G w_n$ is a **derivation** in *G*.
   c. Let $\Rightarrow_G^*$ be the reflexive, transitive closure of $\Rightarrow_G$.
   d. Then the **language generated by *G***, denoted $L(G)$, is $\{w \in \Sigma^* : S \Rightarrow_G^* w\}$.
   e. A language *L* is ***context-free*** if there is some context-free grammar *G* such that $L = L(G)$.

3. In a regular grammar, every rule (production) in *R* must have a right-hand side that is
   a) ε, or
   b) a single terminal symbol, or
   c) a single terminal followed by a single nonterminal.

$S \to bS$, $S \to aT$ $T \to aS$, $T \to b$, $T \to \varepsilon$

4. L is a regular language if and only if L = L(G) for some regular grammar G.
   Construction for of one direction (regular grammar $\to$ FSM) : *Details on slide*
   a. Do it for the example above.

5. Recursive grammar contains rules like $X \to w_1 Y w_2$, where $Y \Rightarrow^* w_3 X w_4$ for some w_1, w_2, w_3, and w_4 in V^*.
6. Self-embedding grammar contains rules like $X \to w_1 Y w_2$, where $Y \Rightarrow^* w_3 X w_4$ and both $w_1 w_3$ and $w_2 w_4$ are in Σ^+.

7. If a CFG G is not self-embedding , the L(G) is ______________

8. Consider our grammar for Bal: **S → (S) | ε | SS** Draw a *derivation tree* for the string (( )) (( ) ( ))

9. Hints for designing context-free grammars

Union of two sets:	$A \to B \mid C$
Concatenation:	$A \to BC$
Generate outside in:	$A \to aAb$

a) $L = \{a^n b^n c^m : n, m \geq 0\}$

b) $L = \{ a^{n_1} b^{n_1} a^{n_2} b^{n_2} \ldots a^{n_k} b^{n_k} : k \geq 0 \wedge \forall i \leq k\ (n_i \geq 0) \}$

c) $L = \{a^n b^m : n \neq m\}$

d) $L = \{w \in \{a, b\}^* : \#_a(w) = \#_b(w)\}$

10. Prove that a grammar is correct: L: $A^n B^n = \{a^n b^n : n \geq 0\}$ G: $S \to a\,S\,b,\ S \to \varepsilon$

a) Show that if $w \in L$, then $S \Rightarrow^* w$.
Induction on what? What to prove by induction? Use this to prove what we want.

b) Show that if $S \Rightarrow^* w$, then $w \in L$.
Induction on what? What to prove by induction? Use this to prove what we want.