Fundamentals

1. *Definitions.* Be able to define carefully:
   - DFA, NFA, and regular expression.
   - PDA and context-free grammar.
   - Ambiguous CFG
   - Chomsky normal form.
   - Turing machine
   - Turing-recognizable set, Decidable set
2. *Theorems.* The following are the cornerstone results of the course. Be able to state each of them correctly.
   - Equivalence theorems between NFAs, DFAs, and regular expressions.
   - Equivalence theorems between PDAs and context-free grammars.
   - The Pumping Lemmas (regular and context-free)
   - The undecidability of the halting problem.
3. *Proofs.* You should be able to provide proofs for the following fundamental results.
   - For each class of languages (regular, context-free, etc) know which of the basic closure properties (closure under union, intersection, complement, etc) holds; be able to prove or provide counterexamples where appropriate.
   - Proof of the Pumping Lemmas for regular sets.
   - Simple proofs of non-regularity and non-context-freeness for various languages, *using* each of the pumping lemmas;
   - Proof of the undecidability of the halting problem.

Sample questions:

1. Be able to translate between NFAs, DFA, and regular expressions.
2. Be able to decide whether equations between regular expressions are valid.
3. Be able to put a context-free grammar into Chomsky normal form.

4. Complete each of the following sentences with one of the following: "regular", "context-free", "decidable", "recognizable", "co-recognizable", or "none of the above". You are to make the strongest claim that is always true. For example, for the question "If A and B are regular then $A \cup B$ is ..." completing the sentence with "decidable" would make a true statement, but would earn you no points here, since completing the sentence with "regular" would make a stronger true statement. (Recall that to say that a set X is co-recognizable is simply to say that the complement of X is recognizable.)

   - If A is context-free and B is context-free then $A \cup B$ is ....
   - If A is context-free and B is context-free then $A \cap B$ is ....
   - If A is context-free and B is context-free then $A \setminus B$ is ....
   - If A is context-free and B is context-free AB is ....
   - If A is context-free then $\Sigma^* \setminus A$ is ....
   - If A is context-free then A^* is ....

   NOW, replace "context-free" by "regular", "decidable", "recognizable", or "co-recognizable". Many possible questions here. Mix and match. Lots of fun.

5. Categorize each of the following languages as being regular, context-free, or "neither." Give the most restrictive category that applies.

   (a) $\{a^n b^{2m} \mid n \geq 0, m \geq 0\}$
   (b) $\{a^n b^m \mid n \neq m\}$
   (c) $\{a^n b^m c^k d^l \mid 2n = 3k \text{ or } 5m = 7l\}$
   (d) $\{a^n b^m c^k d^l \mid 2n = 3k \text{ and } 5m = 7l\}$
   (e) $\{a^n b^m c^k d^l \mid 2n = 3m \text{ and } 5k = 7l\}$
   (f) $\{a^n b^m c^k d^l \mid 2n = 3l \text{ and } 5k = 7m\}$
   (g) $\{x \mid \#(a) \text{ in } x > \#(b) \text{ in } x\}$
   (h) $\{x \mid \#(a) \text{ in } x < 2\#(b) \text{ in } x\}$
   (i) $\{a^n \mid n \text{ is a power of } 2\}$

6. Categorize each of the following languages as being Turing-decidable, Turing-recognizable, co-Turing-recognizable, or "none of these". Give the most restrictive category that applies.

   (a) $\{< G > \mid G \text{ is a CFG and } L(G) = \emptyset\}$
   (b) $\{< G, H > \mid G \text{ and } H \text{ are CFGs and } L(G) = L(H)\}$
   (c) $\{< G, H > \mid G \text{ and } H \text{ are CFGs and } L(G) \subseteq L(H)\}$

(d) $\{< M > \mid M$ is a Turing machine with fewer than 117 states.$\}$

(e) $\{< M > \mid M$ is a TM accepting all inputs$\}$

(f) $\{< M > \mid M$ is a TM taking fewer than 117 steps on some input.$\}$

(g) $\{< M > \mid M$ is a TM taking fewer than 117 steps on all inputs.$\}$

(h) $\{< M > \mid M$ is a TM taking fewer than 117 steps on at least 117 different inputs.$\}$

(i) $\{< M > \mid M$ is a TM accepting at least 117 different inputs.$\}$

(j) $\{< M > \mid M$ is a TM accepting at most 117 different inputs.$\}$

(k) $\{< M > \mid M$ is a TM accepting all inputs of length at most 117.$\}$

(l) $\{< M, N > \mid M$ and N are TMs and $L(M) = L(N)$.$\}$

(m) $\{< M, N > \mid M$ and N are TMs and $L(M) \neq L(N)$.$\}$

(n) $\{< P, Q > \mid P$ and Q are C++ programs and P and Q compute the same function.$\}$

(o) $\{< P, Q > \mid P$ and Q are C++ programs and P and Q compute different functions.$\}$

(p) $\{< P > \mid P$ is a C++ program and P terminates normally on input 0.$\}$

(q) $\{< P > \mid P$ is a C++ program and P terminates normally on all inputs.$\}$