This exam begins for you when you open (or peek inside) this packet. It ends at 9:40 AM on Friday 2013 October 11. Between those two times, you may work on the exam as much as you like. Although I do not intend the exam to require more than a couple of hours, you should get started early, in case you want to spend more time. The exam is open-book and open-note:

- You may use all of this course's materials: the Sipser textbook, your class notes, your old homework, and the materials on our course web site. If you missed a class and want to get some other student's notes, then do so before either of you begins the exam. You may not share any materials with any other person while you are taking the exam.

- You may cite theorems and examples from class, and from the assigned sections of our textbook. You do not have to reprove them. On the other hand, you may not cite results that we have not studied.

- You may not consult any other books, papers, Internet sites, etc. You may use a computer for viewing the course web site, typing up your answers, and e-mailing with me. If you want to use a computer for other purposes, then confirm with me first.

- You may not discuss the exam in any way — spoken, written, etc. — with anyone but me, until everyone has handed in the exam. During the exam period you will inevitably see your classmates around campus. Refrain from asking even seemingly innocuous questions such as "Have you started the exam yet?" If a statement or question conveys any information about the exam, then it is not allowed. If it conveys no information, then you have no reason to make it.

Feel free to ask clarifying questions in person or over e-mail. You should certainly ask for clarification if you believe that a problem is misstated. Check your e-mail occasionally, in case I send out a correction.

Your solutions should be thorough, self-explanatory, neat, concise, and polished. You might want to work on scratch paper, and then recopy your solutions. Alternatively, you might want to type your solutions. Always show enough work and justification so that a typical classmate could understand your solutions. If you cannot solve a problem, then write a brief summary of the approaches you've tried. Partial credit is often awarded. Present your solutions in the order assigned, in a single stapled packet.

Good luck.

A. Give three languages A, B, C over $\Sigma = \{0, 1\}$ such that $A \subseteq B \subseteq C$, A is regular, B is not regular, and C is regular.

B. Compared to most programming languages, Python has clean, simple syntax. However, even its syntax is not truly context-free. We will essentially prove so in this problem. To keep things manageable, we will work on a simplified version of the language, called Baby Python.

Let `\n` be the new-line character, which ends each line of Baby Python code. Let `\t` be the tab character and `\s` the space character. Let R be a regular expression that generates all other single characters — for example, `a`, `+`, etc. A string is a valid line of Baby Python code if and only if it is matched by the regular expression

$$(\texttt{\textbackslash t} \cup \texttt{\textbackslash s})^*(R(R \cup \texttt{\textbackslash t} \cup \texttt{\textbackslash s})^* \cup \epsilon)\texttt{\textbackslash n}.$$

The substring generated by the initial $(\texttt{\textbackslash t} \cup \texttt{\textbackslash s})^*$ is called the *indentation* of the line. The unusual feature of Baby Python (and Python) is its white-space sensitivity. After an `if`, `for`, `while` etc. statement comes a *block* consisting of one or more lines, all of which must have the same indentation. For example, here is an `if` statement followed by a three-line block, in which all three lines are indented four spaces (not the same thing as a four-space tab, as many a CS 111 student has learned the hard way), followed by a last statement outside the block.

```
if miscreantFound:
    alert(detective, ‘‘We got your man.’’)
    bonus += 10000
    purchaseSnowmobile()
water(housePlants)
```

(In Python, blocks can be nested. In Baby Python they cannot.) Let A be the language consisting of all valid blocks in Baby Python. Prove that A is not context-free.

C. Suppose that we convert a CFG G to a PDA P using the usual algorithm. Let $w \in L(G) = L(P)$, and let n be the length of w. How tall can P's stack grow, in the process of accepting w? How does your answer change, if G is assumed to be in Chomsky normal form?

D. Let M_1 and M_2 be two DFAs over the same alphabet Σ. Let p_1 and p_2 be their numbers of states, respectively. Suppose that M_1 and M_2 agree on all strings w such that $|w| \leq p_1 p_2$ — meaning that if $|w| \leq p_1 p_2$ then either both M_i accept w or both M_i reject w. Prove that $L(M_1) = L(M_2)$. Hint: Construct a product DFA of some kind.

E. How many hours have you spent on this exam? (Your answer does not affect your score. I will not read your answer until I have graded your other problems.)