

A. Let $A = \{a^m b^n c^n : m \geq 0, n \geq 0\}$. Then A is a CFL, generated by the CFG

$$S \rightarrow DE, \quad D \rightarrow \epsilon | aD, \quad E \rightarrow \epsilon | bEc.$$

Similarly, $B = \{a^n b^n c^m : m \geq 0, n \geq 0\}$ is a CFL. Their intersection is $A \cap B = \{a^n b^n c^n : n \geq 0\}$, which is not a CFL, as we've proved in class.

B. I assume that the code to be scraped is already in the string `programCode`. My code is

```
regexp = r'""([\^"]+)"\s+def\s+([\^:]+):|def\s+([\^:]+):\s+""([\^"]+)"'
re.findall(regexp, programCode)
```

The regular expression consists of two halves stuck together with `|`. The left half handles the old style of documentation, and the right half handles the new style. Let me explain the left half. It searches for a string enclosed in triple quotation marks, followed by white space, followed by `def`, followed by white space, followed by a function header, followed by `:`. It contains two groups, for capturing the comment string and the function header. The right half is similar.

[My regular expression is not perfect. For example, it incorrectly captures part of

```
if undef (a, b):
    """About to divide by zero. This comment is pretty useless --- in
    fact, this whole fragment is rather dumb --- but such code could
    exist, in a poorly written Python program."""
    print "error: trying to divide", a, "by zero"
else:
    quotient = a / b
```

Perfecting regular expressions takes more time than is available on an in-class exam. My grading reflects this reality.]

C. This language is not regular. To prove so, assume for the sake of contradiction that A is regular. Let p be a pumping length for A . Let

$$w = a^p b^p c.$$

Then w is a string in A of length at least p . So there exist strings x, y, z such that $w = xyz$, $|xy| \leq p$, $|y| \geq 1$, and $xy^i z \in A$ for all $i \geq 0$. Because $|xy| \leq p$, we know that xy consists entirely of as . Therefore y also consists entirely of as ; namely, $y = a^k$ for some k satisfying $1 \leq k \leq p$. But then $xy^2 z = a^{p+k} b^p c \notin A$, which is a contradiction. We conclude that A is not regular.

D. [Justification is not required. I give it anyway, for educational value.]

1. FALSE. [The class of CFLs is closed under union. If it were also closed under complementation, then it would be closed under intersection (by DeMorgan's law). But the class of CFLs is not closed under intersection (by Problem A).]
2. TRUE. [If $\Sigma^* - A$ is finite, then $\Sigma^* - A$ is regular, and hence A is regular.]
3. TRUE. [See the definition of acceptance of a string by an NFA.]
4. FALSE. [A CFG in CNF is allowed to contain the rule $S \rightarrow \epsilon$. Also, we have shown that all CFGs can be put into CNF. "All" includes the CFGs that are able to derive ϵ .]
5. TRUE. [The transition function takes as input a state, and an alphabet symbol or ϵ . It produces as output a set of states.]
6. TRUE. [Just reverse the right side of each rule, and you get a CFG for the reversed language.]