## Small Problem 5: PCFG Sentence Completion

**Summary**

**Given:**

PCFG Two strings and

**Find:**

What is the conditional probability that the grammar will generate given that the string begins with the prefix string ? See the details file for the grammar and the queries.

**Metrics:**

Square of the difference in negative log probability (“surprise”) between the true and the computed conditional probability.

**Details**

In this problem, you are given a probabilistic context free grammar and a prefix of a string generated by that grammar. Your goal is to evaluate the conditional probability that the grammar will generate the complete string for a given suffix : .

Grammar: Note that this grammar produces strings of unbounded length, but produces a finite string with probability 1.

|  |  |  |  |
| --- | --- | --- | --- |
| S → AB (0.25)  S → BC (0.2)  S → AC (0.4)  S → CA (0.15) | A → a (0.05)  A → b (0.3)  A → S (0.65) | B → b (0.5)  B → c (0.3)  B → d (0.2) | C → d (0.35)  C → e (0.1)  C → S (0.55) |

Query 1: What is the conditional probability that a string beginning with “bd” terminates with the complete string “bdcb”: where is the complete string that is generated by S.

Query 2: What is ?

Metric: Square of the difference in negative log probability (“surprise”) between the true and the computed conditional probability. .