Suppose we have the following PCFG for modeling some simple English sentences.

* Terminals {the, a, some, steak, fork, sauce, Alice, golden, hot, ate, I ...}
* Non-terminals {S, NP, VP, PP, NBAR, NN, PRPN, PRON, P, JJ, VB}

|  |  |  |
| --- | --- | --- |
| 1.0 S → NP VP | 0.6  VP → VB NP  0.3  VP → VB  0.1  VP → VP PP | 1.0 PP → P  NP |
| 1.0 PRON → I |
| 0.5 NP → DT NBAR  0.3 NP → PRPN  0.2 NP → PRON | 0.75  NBAR → NN  0.15  NBAR → NBAR PP  0.1    NBAR → JJ NBAR | 1.0 PRPN → Alice |
| 0.5 JJ → golden  0.5 JJ → hot |
| 0.4 DT → the  0.4 DT → a  0.2 DT → some | 0.35  NN → steak  0.35  NN → sauce  0.3    NN → fork | 1.0 VB → ate |
| 1.0 P → with |

1. What is the probability of the sentence “I ate” in the grammar?
   1. Pr(S → NP VP) \* Pr(NP → PRON | NP) \* Pr (PRON → I | PRON) \* Pr(VP → VB | VP) \* Pr(VB → ate | VB) = 1\* 0.2\* 1 \*0.3 \* 1 = 0.06
2. What would be the probability of the sentence if it had a second parse?