**2. Language Modeling**

**Dataset D:**

<s> I am </s>

<s> am I </s>

<s> am am </s>

**Bigram probabilities:**

|  |  |  |
| --- | --- | --- |
|  | **Model U** | **Model S** |
| **P(I | <s>)** | 1/3 | 2/5 |
| **P(am | <s>)** | 2/3 | 3/5 |
| **P(</s> | <s>)** | 0 | |
| **P(I | I)** | 0/2 = 0 | 1/4 |
| **P(am | I)** | 1/2 | 2/4 = 1/2 |
| **P(</s> | I)** | 1/2 | 2/4 = 1/2 |
| **P(<s> | I)** | 0 | |
| **P(I | am)** | 1/4 | 2/6 = 1/3 |
| **P(am | am)** | 1/4 | 2/6 = 1/3 |
| **P(</s> | am)** | 2/4 = 1/2 | 3/6 = 1/2 |
| **P(<s> | am)** | 0 | |

Key sentence in D with higher probability in Model S than Model U: <s> I am </s>

Model U: PModelU(<s> I am </s>) = PModelU (I | <s>) \* PModelU (am | I) \* PModelU (</s> | am) = 1/3 \* 1/2 \* 1/2 = 1/12

Model S: PModelS(<s> I am </s>) = PModelS (I | <s>) \* PModelS (am | I) \* PModelS (</s> | am) = 2/5 \* 1/2 \* 1/2 = 1/10

Therefore, PModelU(<s> I am </s>) < PModelS(<s> I am </s>).