

Übung 4: Sprachidentifizierer

Rechenübungen

Gegeben seien die folgenden Häufigkeiten von Buchstabenfolgen:

	a	b	</s>
a	14	4	2
b	6	12	2
<s>	1	1	0

$f(a,a) = 14$	$f(b,a) = 6$	$f(<s>,a) = 1$
$f(a,b) = 4$	$f(b,b) = 12$	$f(<s>,b) = 1$
$f(a,<s>) = 2$	$f(b,<s>) = 2$	$f(<s>,<s>) = 0$

Die Häufigkeit von "ab" ist hier 4.

- Berechnen Sie die Wahrscheinlichkeiten eines Markowmodelles 1. Ordnung.
- Welche Wahrscheinlichkeit hat die Buchstabenfolge *ababb*?

11)
$$P(x_2 | x_1) = \frac{f(x_1, x_2)}{\sum_x f(x_1, x)}$$

$$\sum_x f(a, x) = f(a,a) + f(a,b) + f(a,<s>) = 20$$

$$\sum_x f(b, x) = 6 + 12 + 2 = 20$$

'ba'

$$P(a|b) = \frac{f(b,a)}{\sum_x f(b,x)} = \frac{6}{20} = 0.3$$

$$\sum_x f(<s>, x) = 2$$

$$P(a|a) = 0.7$$

$$P(b|a) = 0.2$$

$$P(<s>|a) = 0.1$$

$$P(b|b) = \frac{12}{20} = 0.6$$

$$P(a|<s>) = 0.5$$

$$P(b|<s>) = 0.5$$

$$P(<s>|<s>) = 0$$

$$P(<s>|b) = 0.1$$

12)
$$P(ababb) = P(a,b,a,b,b) = P(a|<s>) \cdot P(b|a) \cdot P(a|b) \cdot P(b|a) \cdot P(b|b) \cdot P(<s>|b)$$

$$= 0.5 \times 0.2 \times 0.3 \times 0.2 \times 0.6 \times 0.1$$

$$= 0.0036$$