

W modelu IBM 1 zakładamy że prawdopodobieństwo zdania w języku obcym F o długości l_F i dopasowania a pod warunkiem zdania w naszym języku N o długości l_N wynosi:

$$P(F, a|N) = \frac{\epsilon}{(1 + l_N)^{l_F}} \prod_{j=1}^{l_F} P(F_j|N_{a(j)})$$

gdzie $P(F_j|N_{a(j)})$ jest prawdopodobieństwem że słowo F_j (j -te słowo F jest tłumaczeniem $N_{a(j)}$ (słowa o numerze $a(j)$ w N) zaś ϵ (zależne od l_F i N , ale niezależne od a i F) jest dobrane tak by prawdopodobieństwa sumowały się do 1.

Uwaga: Do zdania N na pozycji 0 dodajemy fikcyjne słowo, dzięki temu każde słowo z F ma odpowiednik w N (choć nie każde słowo z N ma odpowiednik w F). Ponadto słowo z N może odpowiadać wielu słowom z F . Oznacza to że a jest dowolną funkcją z $[1, \dots, L_F]$ w $[0, \dots, L_N]$.

1. Sprawdź że w modelu IBM 1 zachodzi równość:

$$\sum_{F,a} P(F, a|N) = \epsilon$$

gdzie suma przebiega po wszystkich zdaniach długości L_F i w wszystkich funkcjach $a : [1, \dots, L_F] \mapsto [0, \dots, L_N]$. Innymi słowy $\epsilon = P(L_F|N)$.

2. Spróbuj ustawić we właściwej kolejności następujące ciągi słów angielskich (po jednym w linii):

```
best by explained failure human is stupidity usually
infidels hates he
calculations useful is symbolic for it
Joe Eve loves
```