

Aufgabe.1:

Wir bekommen immer dieselbe falsche Übersetzung für jedes Satz.

Aufgabe.2:

Diese TMX Datei speichert Übersetzungsdatei und dient zum Datenaustausch zwischen unterschiedliche Programme.

Um die Übersetzungsrichtungen zu ändern, kann man den Tag von Sprachen Z.B. <srclang> und <tuv lang='DE-DE'> verändern.

Es gibt schon über 300 Segmentationsregeln für Deutsch, im Vergleich zu den 22 englischen Unterstützungen, die deutsche sind viel mehr detailliert und umfassend.

Aufgabe.3

Die Bigramme von den Trainingskorpus:

$e^{(i-1)} \backslash e^i$	START	the	house	is	small	little	building	home	END
START									
the	4								
House		2							
Is			2				2		
Small				3					
Little				1					
Building		1						1	
home		1							
END					3	1			

$P(e|e') = \text{Count}(e', e) / \text{Count}(e)$, deshalb kriegen wir folgende Parameter:

$P(\text{the}|\text{START})=1$, $P(\text{house}|\text{the})=0.5$, $P(\text{is}|\text{house})=1$, $P(\text{small}|\text{is})=0.75$, $P(\text{END}|\text{small})=1$,
 $P(\text{building}|\text{the})=0.25$, $P(\text{is}|\text{building})=1$, $P(\text{home}|\text{the})=0.25$, $P(\text{building}|\text{home})=1$, $P(\text{little}|\text{is})=0.25$,
 $P(\text{END}|\text{little})=1$.

$f = \text{'das Haus is klein'}$, $e_1 = \text{'the building is little'}$, $e_2 = \text{'the house is small'}$.

1. $p(e_1, f)$:

$$P(e_1) = P(\text{the}|\text{START}) * P(\text{building}|\text{the}) * P(\text{is}|\text{building}) * P(\text{little}|\text{is}) * P(\text{END}|\text{little})$$

$$= 1 * 0.25 * 1 * 0.25 * 1 = 0.0625$$

$$p(f|e_1) = P(\text{das}|\text{the}) * P(\text{Haus}|\text{building}) * P(\text{ist}|\text{is}) * P(\text{klein}|\text{little}) / (4+1)^4$$

$$= 0.4 * 0.3 * 0.7 * 0.7 / 625 = 0.00009408$$

$$\text{Also } p(e_1, f) = p(e_1) * p(f|e_1) = 0.00000588.$$

2. $p(e_2, f)$:

$$P(e_2) = P(\text{the}|\text{START}) * P(\text{house}|\text{the}) * P(\text{is}|\text{house}) * P(\text{small}|\text{is}) * P(\text{END}|\text{small})$$

$$= 1 * 0.5 * 1 * 0.75 * 1 = 0.375$$

$$p(f|e_2) = P(\text{das}|\text{the}) * P(\text{Haus}|\text{house}) * P(\text{ist}|\text{is}) * P(\text{klein}|\text{small}) / (4+1)^4$$

$$= 0.4 * 0.5 * 0.7 * 0.8 / 625 = 0.0001792$$

$$\text{Also } p(e_2, f) = p(e_2) * p(f|e_2) = 0.0000672$$