

12. Übungsblatt - Abgabe: 03.02.2015

Aufgabe 12.1 - Maschinelle Übersetzung

Gegeben sind die folgenden Übersetzungswahrscheinlichkeiten für einzelne Wörter.

Mann	man	0.7
Buch	book	0.9
liest	reads	0.5
der	the	0.7
ein	a	0.8

- Berechnen Sie die Übersetzungswahrscheinlichkeit für die beiden Satzpaare *The man reads a book* - *Der Mann liest ein Buch* und *The man reads a book* - *Ein Buch liest der Mann*. Nehmen Sie an, dass eine Wortalignierung vorliegt, die die sich jeweils semantisch entsprechenden Wörter aligniert. Nehmen Sie das Übersetzungsmodell, das Umordnungskosten mit einbezieht.
- Was würde bei dem Modell ohne Umordnungskosten passieren?
- Wenn wir die wahrscheinlichste deutsche Übersetzung d für einen englischen Satz e finden wollen, berechnen wir $\max_d P(d|e)$ als $\max_d P(e|d) * P(d)$. Übersetzungswahrscheinlichkeiten für einzelne Wörter stehen uns aber in beide Richtungen zur Verfügung. Erläutern Sie an einem Beispiel, warum es trotzdem sinnvoll ist, die zweite, scheinbar kompliziertere Variante der Formel zu benutzen.

Aufgabe 12.2 - google translate

Übersetzungssysteme stehen vor dem Problem lexikalische Ambiguitäten auflösen zu müssen. Gehen Sie auf <https://translate.google.de>. Betrachten Sie die lexikalisch ambigen Wörter *Schloss*, *Strauß* und *Kiefer* mit jeweils zwei Hauptlesarten.

- Wie übersetzt google translate die Begriffe ohn irgendeinen Kontext. Vermuten Sie wieso.
- Versuchen Sie jetzt für die jeweils andere Lesart je einen deutschen Satz zu finden, in dem für einen Mensch diese Lesart offensichtlich ist und das System den Begriff richtig übersetzt.
- Versuchen Sie je einen Satz für die Lesart aus b zu finden, in dem die Lesart wieder für einen Menschen klar ist, aber das System falsch übersetzt. Beschreiben Sie kurz Ihr Vorgehen.

Abgabe in Gruppen von bis zu drei Studierenden am **03.02.2015** vor der Vorlesung.