25. Januar 2013

GKI - Hausaufgaben 5

Tao Xu, 343390 - Mitja Richter, 324680 - Björn Kapelle, 320438 - Marcus Weber, 320402

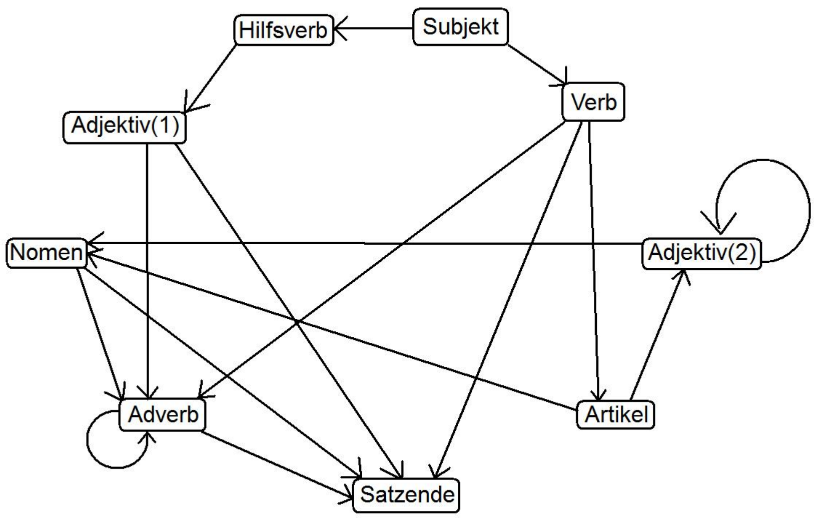
Aufgabe 1

1.a)

1.b)

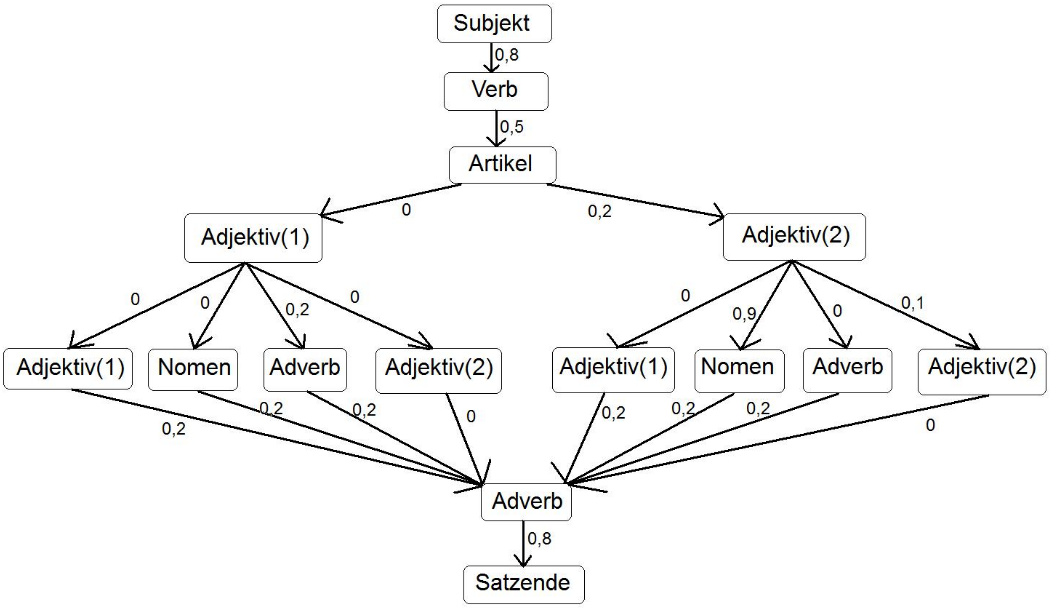
Aufgabe 2

2.a)



2.b)

„He"∈ {Subjekt}
„shoots"∈ {Verb}
„the"∈ {Artikel}
„unwell"∈ {Adjektiv(1), Adjektiv(2)}
„well"∈ {Adjektiv(1), Nomen, Adverb, Adjektiv(2)}
„badly"∈ {Adverb}
Daraus folgt das folgende Bild.



Die einzigen möglichen Folgen von Wortarten beschreiben Pfade von der Wurzel (Subjekt) zum Blatt (Satzende) bei dem jede Kante eine positive Wahrscheinlichkeit hat. Es gibt nur einen Pfad, der das erfüllt: Subjekt-Verb-Artikel-Adjektiv(2)-Nomen-Adverb-Satzende.
Insbesondere gibt es damit eine Möglichkeit diesen Satz zu bilden.
Im weiteren verwenden wir Abkürzungen:
Sub - Subjekt
Ver - Verb
Hil - Hilfsverb
Art - Artikel
Adj(2) - Adjektiv(2)
Nom - Nomen
Adv - Adverb
SE - Satzende

$P(\text{passende Wortartenfolge}) = P(x_0 = Sub, x_1 = Ver, x_2 = Art, x_3 = Adj(2), x_4 = Nom, x_5 = Adv, x_6 = SE)$
$= P(Ver|Sub)P(Art|Ver)P(Adj(2)|Art)P(Nom|Adj(2))P(Adv|Nom)P(SE|Adv)$
$= 0,8 \cdot 0,5 \cdot 0,2 \cdot 0,9 \cdot 0,2 \cdot 0,8$
$= 0,1152$

Das ist allerdings nur die Wahrscheinlichkeit für die passende Folge von Wortarten. Diese muss nun mit den Einzelwahrscheinlichkeiten für die Worte noch multipliziert werden.
$P($ „He shoots the unwell well badly.“$)$
$= P(passendeWortartenfolge) \cdot P(He|Sub) \cdot P(shoots|Ver) \cdot P(the|Art) \cdot P(unwell|Adj(2)) \cdot P(well|Nom) \cdot P(badly|Adv)$
$= 0,1152 \cdot 0,5 \cdot 0,5 \cdot 0,6 \cdot 0,5 \cdot 0,4 \cdot 0,5 = 0,0001728 = 0,01728\%$

Die Wahrscheinlichkeit für den Satz „He shoots the unwell well badly.“beträgt 0,01728 %.

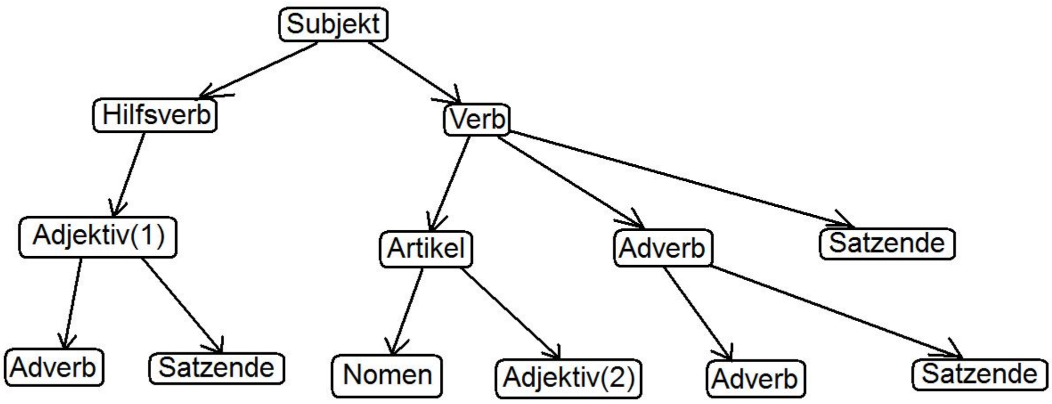
2.c)

$y_0 = She \Rightarrow x_0 \in \{Subjekt\}$
$y_1 = is \Rightarrow x_1 \in \{Verb, Hilfsverb\}$
$\Rightarrow x_2 \in \{Artikel, Adverb, Satzende, Adjektiv(1)\}$

$x_4 = Satzende \Rightarrow x_3 \in \{Verb, Adjektiv(1), Nomen, Adverb\}$
$y_3 = well \Rightarrow x_3 \in \{Adjektiv(1), Nomen, Adverb, Adjektiv(2)\}$
$\Rightarrow x_3 \in \{Adjektiv(1), Nomen, Adverb\}$
$\Rightarrow x_2 \in \{Hilfsverb, Artikel, Adjektiv(2), Verb, Adjektiv(1), Nomen, Adverb\}$

Aus diesen beiden Restriktionen für x_2 kann man folgern, dass: $x_2 \in \{Artikel, Adverb, Adjektiv(1)\}$

2.d)



Sätze mit genau 3 Wörtern entsprechen Pfaden von der Wurzel (Subjekt) zu Blättern, die mit „Satzende"bezeichnet sind und diese Pfade sollen über 3 Kanten gehen.
Hierfür gibt es folgende geordnete Wortartkombinationen:
(i) Subjekt-Hilfsverb-Adjektiv(1)-Satzende (ii) Subjekt-Verb-Adverb-Satzende
$P($ Satz mit genau 3 Wörtern $) = P(x_0 = Sub, x_1 = Hil, x_2 = Adj(1), x_3 = SE) + P(x_0 = Sub, x_1 = Ver, x_2 = Adv, x_3 = SE)$
$= P(Hil|Sub) \cdot P(Adj(1)|Hil) \cdot P(SE|Adj(1)) + P(Ver|Sub) \cdot P(Adv)|Ver) \cdot P(SE|Adv)$
$= 0,2 \cdot 1 \cdot 0,8 + 0,8 \cdot 0,2 \cdot 0,8$
$= 0,16 + 0,128 = 0,288$

Die Wahrscheinlichkeit für einen Satz mit genau 3 Wörtern beträgt 28,8 %.

$M((i))$ bezeichne die Menge der Möglichkeiten für den Fall (i) und $|M(i)|$ bezeichne die Anzahl an Möglichkeiten für den Fall (i), , für (ii) analog.
Mit $|Wortart|$ bezeichnen wir die Anzahl möglicher Worte aus der Wortart.
$|M(i)| = |Subjekt| \cdot |Hilfsverb| \cdot |Adjektiv(1)| = 2^3 = 8$
$|M(ii)| = |Subjekt| \cdot |Verb| \cdot |Adverb| = 2^3 = 8$
Die Gesamtanzahl an Möglichkeit ergibt sich aus: $|M((i))| + |M((ii))| - |M((i)) \cap M((ii))|$ Dazu betrachte: $Subjekt \cap Subjekt = \{He, She\}$
$Hilfsverb \cap Verb = \{is\}$
$Adjektiv(1) \cap Adverb = \{well\}$
$\Rightarrow |M((i)) \cap M((ii))| = |Subjekt \cap Subjekt| \cdot |Hilfsverb \cap Verb| \cdot |Adjektiv(1) \cap Adverb| = 2$
$\Rightarrow$ Anzahl verschiedener Möglichkeiten für einen Satz mit genau 3 Wörtern $= 8 + 8 - 2 = 14$