

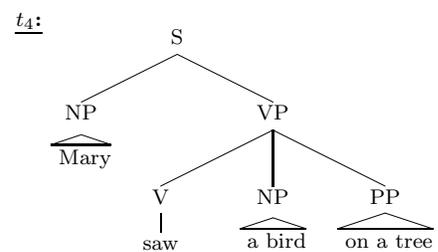
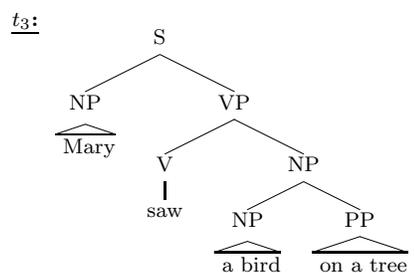
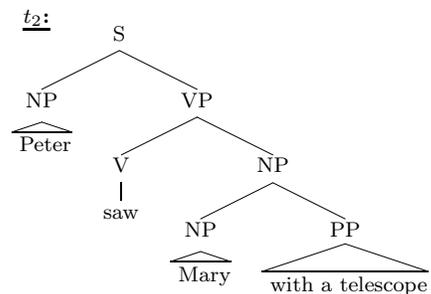
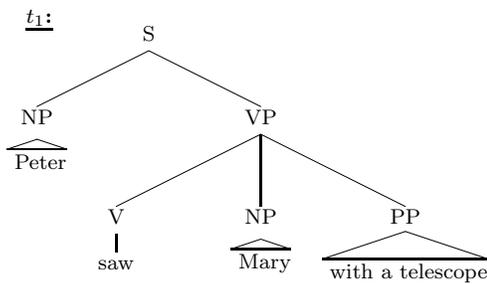
Statistisches Parsing

Detlef Prescher

April 28, 2007

Hausaufgabe 2

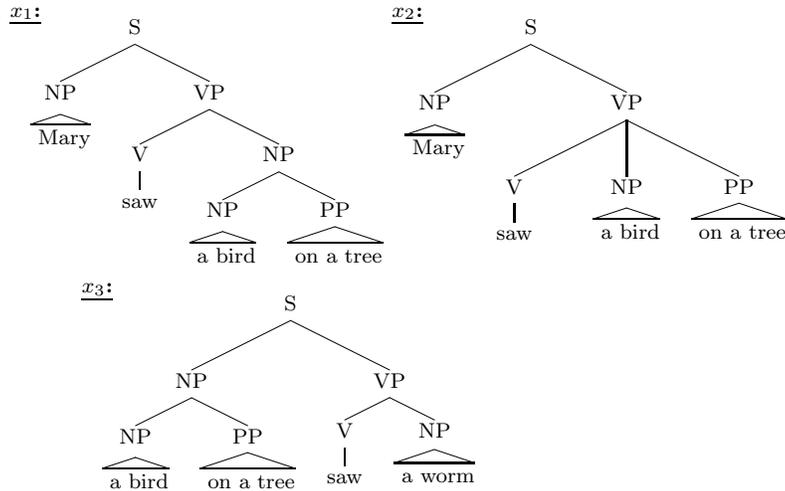
(i) Gib eine möglichst kleine kontextfreie Grammatik an, die die folgenden Syntaxbäume generiert:



Gib eine probabilistische Version dieser Grammatik an, sodass alle Regeln mit derselben linken Seite die gleiche Wahrscheinlichkeit haben. Berechne damit die Wahrscheinlichkeiten der Bäume t_1, t_2, t_3 und t_4 und diskutiere das Ergebnis.

(ii) Lexikalisiere in der Grammatik unter (i) alle PP-Regeln. Probabilisiere die resultierende Grammatik so, dass eine bessere Disambiguierung erfolgt.

(iii) Gib eine möglichst kleine kontextfreie Grammatik an, die die folgenden Syntaxbäume generiert:



Gib eine probabilistische Version dieser Grammatik an, sodass alle Regeln mit derselben linken Seite die gleiche Wahrscheinlichkeit haben. Berechne damit die Wahrscheinlichkeiten der Bäume x_1, x_2 und x_3 und diskutiere das Ergebnis.

(iv) Angenommen, Du bekommst die Grammatik in (iii) vorgelegt und einen Korpus in dem der Satz "a bird on a tree saw a worm" sehr viel häufiger vorkommt als der Satz "Mary saw a bird on a tree". Welchen Grammatikregeln würdest Du dann eine sehr hohe Wahrscheinlichkeit zuweisen wollen? Gib ein fiktives Beispiel an. Berechne damit die Wahrscheinlichkeiten der Bäume x_1, x_2 und x_3 und diskutiere das Ergebnis.

(v) Konstruiere eine Grammatik, die die Standard-Nebenbedingung für probabilistische kontextfreie Grammatiken verletzt, die aber doch allen ihren Syntaxbäumen die Wahrscheinlichkeit 1 zuweist!

(**Standard-Nebenbedingung:** die Summe der Wahrscheinlichkeiten aller Regeln mit derselben linken Seite ist immer 1.)

Abgabetermin: Montag, 7. Mai

Besprechung: Donnerstag, 10. Mai