

# Übung zur Vorlesung “Einführung in die Computerlinguistik und Sprachtechnologie”

Wintersemester 2017/2018, Prof. Dr. Udo Hahn, Sven Büchel

Übungsblatt 10 vom 19.01.2018

Abgabe bis 23.01.2018, 23.59 Uhr; per Email (PDF-Format) an sven.buechel@uni-jena.de

## Aufgabe 1 Sprachen zu Automaten

5

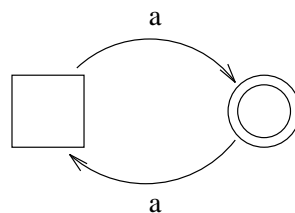
Geben Sie zu jedem der folgenden endlichen Automaten das von ihm verwendete *Alphabet* sowie die von ihm *akzeptierte Sprache* an (also die Menge aller Ketten/Wörter, die der Automat akzeptiert). Verwenden Sie dabei entweder Potenzschreibweise oder zählen Sie die Sprache vollständig auf.

Anmerkungen:

- Ein Kreis in einem Zustandssymbol - also Doppelkreis oder Kreis in einem Quadrat - markiert einen Endzustand.
- In der Vorlesung bestanden die “einzelnen Zeichen” teilweise bereits aus linguistischen Einheiten wie Wörtern (“hospital”) oder Morphemen (“ize”). In diesem Übungsblatt deckt sich das Konzept von einzelnen Zeichen der Automaten mit denen des lateinischen Alphabets.
- Ein Komma zwischen zwei Symbolen auf einer Kante (“Pfeil” zwischen zwei Zuständen) bedeutet, dass für jedes dieser Symbole der Übergang möglich ist. Es ist also eine Kurzform dafür, für jedes der Symbole einen eigenen Pfeil zu zeichnen.

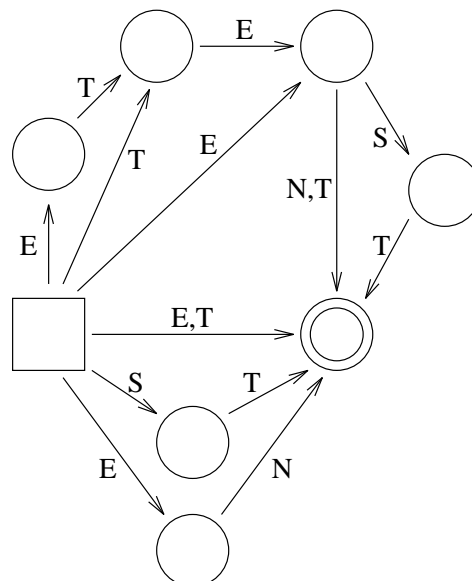
a)

2,5



b)

2,5



## Aufgabe 2 : Parsing (5 Punkte)

In dieser Aufgabe sollen Sie zwei unterschiedliche Parsingverfahren am Satz *Fly to Detroit* vorführen. Verwenden Sie

dafür die folgende Grammatik:

$G = (N, T, P, S)$  mit  
 $N = \{ NP, VP, PP, det, n, v, p \}$   
 $T = \{ fly, to, Detroit \}$   
 $P = \{ S \rightarrow VP PP,$   
     $NP \rightarrow n,$   
     $VP \rightarrow v PP,$   
     $VP \rightarrow v,$   
     $PP \rightarrow p NP,$   
     $n \rightarrow Detroit,$   
     $p \rightarrow to,$   
     $v \rightarrow Fly \}$

Führen Sie einen Top-Down Parse (mit Tiefensuche) und einen Bottom-Up Parse für den Beispielsatz durch. Achten Sie dabei auf die Vorgaben aus dem Parsing-Merkblatt (<http://julielab.de/Students/Merkbl%C3%A4tter.html>).