

Einführung in die Computerlinguistik – deterministische und nichtdeterministische endliche Automaten

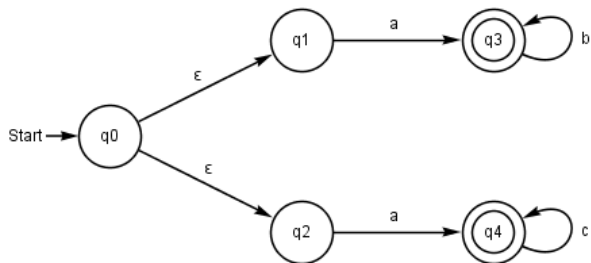
Dozentin: Wiebke Petersen

Foliensatz 4

Äquivalenz von endlichen Automaten

Zwei endliche Automaten heißen **äquivalent**, wenn sie dieselbe Sprache akzeptieren.

endliche Automaten mit ϵ -Übergängen



Zu jedem endlichen Automaten mit ϵ -Übergängen gibt es einen äquivalenten endlichen Automaten ohne ϵ -Übergänge.

ε -Übergänge entfernen

Die ε -Kanten werden Kante für Kante entfernt, wobei folgende Anweisungen zu befolgen sind:

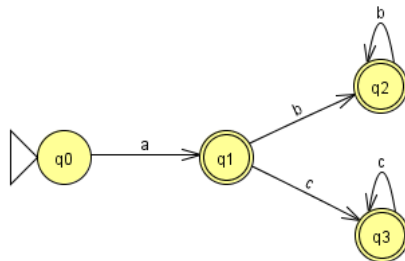
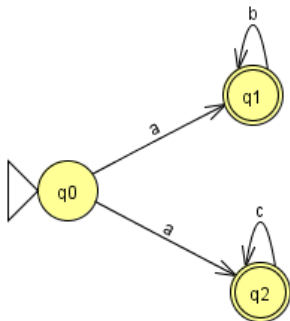
- Wenn die Kante zu einem Endzustand führt, dann wird der Zustand, an dem die Kante beginnt, zu einem Endzustand.

ε -Übergänge entfernen

Die ε -Kanten werden Kante für Kante entfernt, wobei folgende Anweisungen zu befolgen sind:

- Wenn die Kante zu einem Endzustand führt, dann wird der Zustand, an dem die Kante beginnt, zu einem Endzustand.
- Man zeichnet alle Kanten ein, die benötigt werden, damit sich die von dem Automaten akzeptierte Sprache nicht ändert, wenn man die ε -Kante wegläßt:
 - Hierzu überlegt man sich, zu welchen Zuständen man von dem Ausgangszustand der ε -Kante aus hinkommen kann, wenn man genau ein Symbol liest und die ε -Kante zuerst "benutzt". Zu jedem erreichbaren Zustand zeichnet man eine neue Kante ein.
 - Bei mehreren "in Reihe geschalteten" ε -Kanten beginnt man mit der letzten.
 - Bei "im Kreis geschalteten" ε -Kanten werden die betroffenen Zustände miteinander verschmolzen.

deterministische und nichtdeterministische endliche Automaten



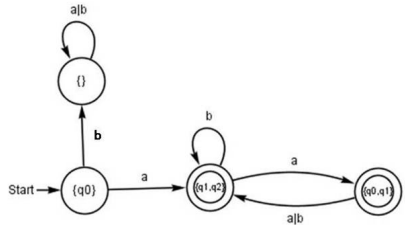
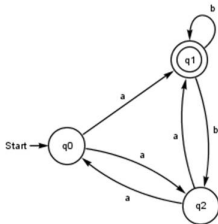
Zu jedem endlichen nichtdeterministischen Automaten gibt es einen endlichen deterministischen Automaten, der dieselbe Sprache akzeptiert.

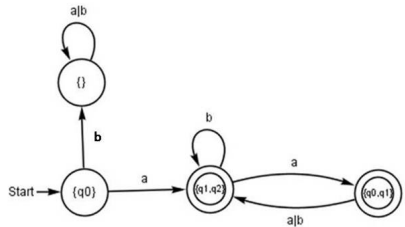
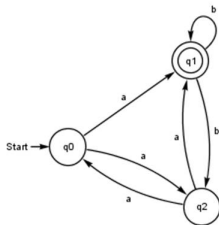
Bemerkung: Jeder Automat mit ϵ -Übergängen ist nichtdeterministisch (überlegen Sie sich, warum das gilt).

Konstruktion eines deterministischen endlichen Automaten (DEA) aus einem nichtdeterministischen (NDEA)

Die Zustände des DEA sind Teilmengen der Zustandsmenge des NDEA:

- Der DEA simuliert die parallele, nichtdeterministische Verarbeitung von Ketten im NDEA.
- Wenn q_0 der Startzustand des NDEA ist, dann ist $\{q_0\}$ der Startzustand des DEA.
- die Eigenschaft Endzustand zu sein "vererbt" sich von dem NDEA auf den DEA. (Wenn q_i ein Endzustand des NDEA ist, dann ist jede Zustandsmenge die q_i enthält, ein Endzustand des DEA).
- für jeden Ausgangszustand und jedes Symbol des Alphabets überlegt man sich, welche Zustände erreicht werden können.

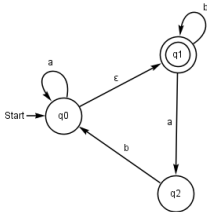




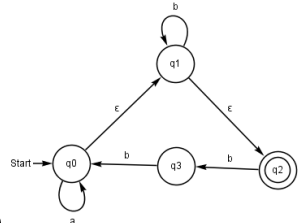
Erklärung an einem Beispiel: in dem rechten deterministischen Automaten steht der Zustand $\{q_1, q_2\}$ dafür, dass man mit der bisher gelesenen Kette (z.B. $abbb$) in dem linken nichtdeterministischen Automaten entweder in dem Zustand q_1 oder in dem Zustand q_2 gelandet ist. Wenn man in q_1 oder in q_2 ist und ein a liest, landet man in q_0 oder q_1 (entspricht Zustand $\{q_0, q_1\}$ im DEA). Wenn man in q_1 oder in q_2 ist und ein b liest, landet man in q_1 oder q_2 (entspricht Zustand $\{q_1, q_2\}$ im DEA). Da q_1 ein Endzustand im NDEA ist, ist $\{q_1, q_2\}$ ein Endzustand im DEA.

Hausaufgaben Teil 1 (BN: Jeweils nur eine der zwei Teilaufgaben)

1) Gegeben seien die folgenden beiden endlichen Automaten:



(I)



(II)

Bitte bearbeiten Sie die folgenden Aufgaben für jeden der beiden Automaten:

- Geben Sie einen regulären Ausdruck an, der die von dem Automaten akzeptierte Sprache beschreibt.
- Geben Sie einen zu dem endlichen Automaten äquivalenten endlichen Automaten ohne ϵ -Übergänge an.
- Geben Sie einen zu dem endlichen Automaten äquivalenten deterministischen endlichen Automaten an.

Hausaufgaben Teil 2 (BN: Jeweils nur eine der zwei Teilaufgaben)

- 2) Zeichnen Sie das Übergangsdiagramm eines endlichen Automaten, der die folgende Sprache akzeptiert. Geben Sie zusätzlich den regulären Ausdruck an, der die Sprache beschreibt:
- a) die Sprache, die aus allen Wörtern über dem Alphabet $\{a, b\}$ besteht, in denen mindestens drei a 's vorkommen und die auf b enden und nicht mit a beginnen.
 - b) die Sprache, die aus allen Wörtern gerader Länge über dem Alphabet $\{a, b\}$ besteht, die mit a beginnen.