Regular Expression -> Finite Automaton

April 2003

Gegeben: Regular Expression z.B.

$$R = b^* a\, (b|a^*)$$

Ziel: Endlichen Automat finden, der genau die Strings akzeptiert, die durch die RE gebildet werden können.

1 Preprocessing

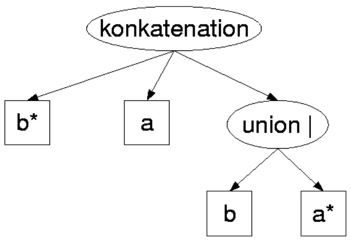
Um den regulären Ausdruck richtig zu interpretieren, ist es wichtig, die Bindungsstärke der einzelnen Operatoren zu kennen. Am stärksten binden (logischerweise) die Klammern, dann folgt der Kleene Star, als nächstes die Konkatenation und am schwächsten bindet die Union.

$$Klammern\ (..) > KleeneStar* > Konkatenation > Union\ |$$

Nun kann der reguläre Ausdruck zuerst in die einzelnen Bestandteile zerlegt werden. Um die Teile klar zu kennzeichnen, können auch Klammern verwendet werden.

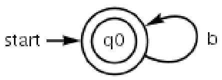
$$R = (b^*)\ (a)\ (b\ |(a^*))$$

Der ganze Ausdruck kann nun auch als Baum dargestellt werden:

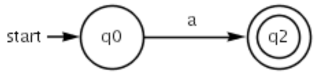


2 Teilausdrücke/Blätter als FAs darstellen

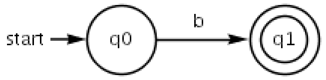
Als nächstes werden die Blätter je als einzelne FAs dargestellt:



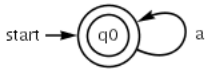
b*:



a:



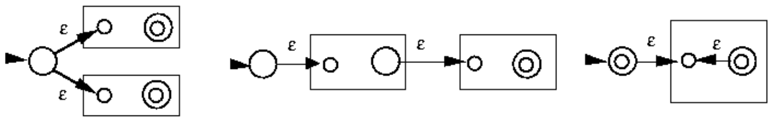
b:



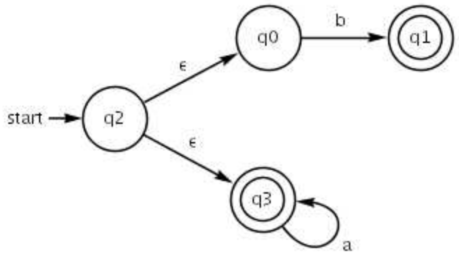
a*:

3 Einzelne FAs miteinander verknüpfen

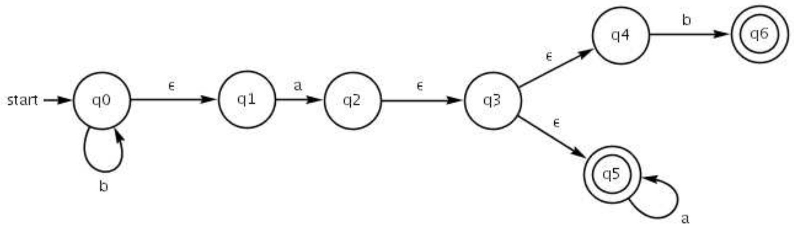
Als letztes müssen die einzelnen Teile miteinander verknüpft werden. Das geschieht bottom-up, immer je nach den Operatoren, die in den inneren Knoten stehen. Das Schema dafür steht im Vorlesungsskript (Kap. 3.5):



Zuerst werden also die zwei Blätter b und a* der Tiefe 2 durch die union zusammengefügt. Dazu wird ein gemeinsamer Startknoten und ε-Uebergänge zu den einzelnen Teilautomaten eingefügt. So wird erreicht, dass der Automat akzeptiert, falls irgendeiner der Teilautomaten akzeptiert.



Nun müssen die drei Knoten b*, a und der neu entstandene Union-Knoten (alle Tiefe 1) durch die Konkatenation zusammengefügt werden. Die einzelnen Automaten werden also hintereinander gehängt und jeweils durch einen ε-Uebergang verbunden. Dabei darf nur der Schluss der Konkatenation akzeptierend bleiben:



Das ist nun ein gültiger NFA mit ε-Uebergängen, der genau diese Sprache akzeptiert.

4 Abschlussarbeiten

Um nun zu einem schönen, kompakten Automaten zu kommen, können die weiteren Schritte angewendet werden. D.h. ε- Uebergänge können noch entfernt werden, der Automat kann noch minimiert werden etc. Das ändert aber nichts an der Korrektheit der ersten Lösung.

Der minimierte Automat (wie ihn der Exorciser als Lösung angibt) sieht so aus: