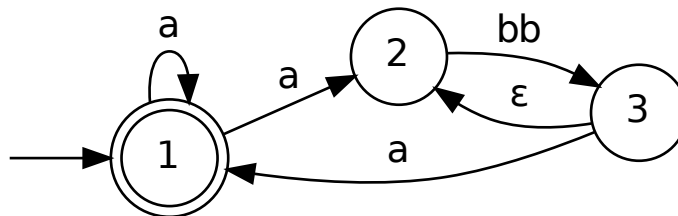


## 8. Übungsblatt (Zusatz) - Abgabe: 07.01.2014

Dieses Übungsblatt ist ein Zusatzübungsblatt. Nur diejenigen von Ihnen, denen wir dies schriftlich mitteilen, müssen das Übungsblatt verpflichtend bearbeiten. (Diejenigen bestätigen uns bitte den Erhalt der entsprechenden Email.) Für alle anderen ist das Blatt freiwillig. Das bedeutet, ihr könnt mit diesem Blatt zusätzliche Punkte bekommen, wenn ihr es komplett bearbeitet, verliert aber auch keinen Freischuss, wenn ihr es nicht bearbeitet und könnt auch kein nicht-abgegebenes Blatt wieder wettmachen.

### Aufgabe 8.1 Automaten

Gegeben sei der folgende NEA über dem Alphabet  $\Sigma = \{a, b\}$ .



- Beschreiben Sie, die von diesem NEA akzeptierte Sprache informell in Worten.
- Geben Sie einen regulären Ausdruck an, der die gleiche Sprache beschreibt wie der Automat.
- Konstruieren Sie nach der Methode der NEA-DEA-Überführung einen zum NEA äquivalenten DEA.

### Aufgabe 8.2 - Evaluation

Bei der Evaluation eines Klassifiers wird das Klassifikationsergebnis für einen Text (Testkorpus) mit einer manuellen Annotation (Goldstandard) verglichen.

Gegeben ist der folgende Testkorpus (als Hilfe für Sie sind die STTS-Tags gegeben):

Das (ART) unbeständige (ADJA) Wetter (NN) der (ART) beiden (PIDAT) letzten (ADJA) Tage (NN) wird (VAFIN) sich (PRF) auch (ADV) am (AP-PR) Montag (NN) fortsetzen (VVINF). (\\$.) Bei (APPR) mäßigem (ADJA) Westwind (NN) treten (VVFIN) vermehrt (ADV) starke (ADJA) Regenschauer (NN) und (KON) gelegentlich (ADV) auch (ADV) Unwetter (NN) auf (PTKVZ). (\\$.) Am (APPR) kommenden (ADJA) Wochenende (NN)

erwarten (VVFİN) wir (PPER) den (ART) kältesten (ADJA) Tag (NN) der (ART) Woche (NN).(\\$.) An (APPR) den (ART) drei (CARD) Tagen (NN) zu (APPR) Wochenanfang (NN) tritt (VVFİN) wieder (ADV) eine (ART) leichte (ADJA) Erwärmung (NN) auf (PTKVZ). (\\$.)

- (a) Annotieren Sie für jedes Wort (einschließlich der Satzzeichen) den Goldstandard (also die tatsächliche Zugehörigkeit zur Klasse ADJA, als + oder -), die 3 Merkmale (entsprechend den Vorlesungsfolien) und das sich daraus ergebende Klassifikationsergebnis (+/-), das das regelbasierte Modell von den Vorlesungsfolien (Statistik I) liefert.
- (b) Geben Sie für diese Evaluation die Konfusionsmatrix an.
- (c) Berechnen Sie aus der Konfusionsmatrix Akkuratheit, sowie Präzision, Recall und F-Score für die Klasse ADJA.

---

Abgabe in Gruppen von bis zu drei Studierenden bis **7.1.2014** 10 Uhr entweder als Email- Anhang im pdf-Format an **i2cl@coli.uni-sb.de** mit dem Betreff **i2cl homework 8** und Ihrer Matrikelnummer oder auf Papier im Briefkasten an der Tür von Raum 1.04 in C7.2.