

## Formale Sprachen: DNA Computing

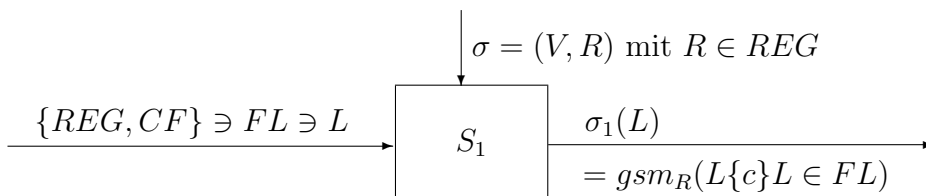
### Splicing mit regulären Regelmengen

Es soll gezeigt werden, dass  $S_1(FL, REG) \subseteq FL$  für  $FL \in \{REG, CF\}$ . Sei dazu  $L \in FL$  eine beliebige Sprache mit  $L \subseteq V^*$  und  $\sigma = (V, R)$  mit  $R \in REG$ . Es wird ein verallgemeinerter endlicher Automat  $gsm_R$  konstruiert, so dass

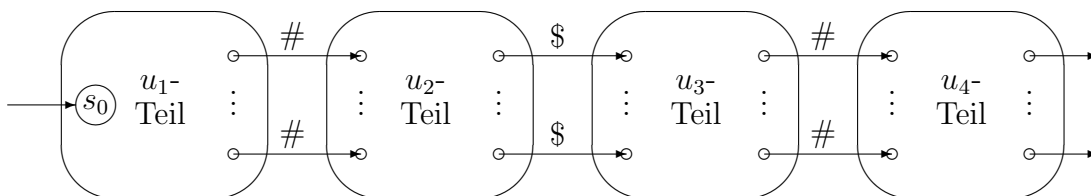
$$gsm_R(L\{c\}L) = \sigma_1(L)$$

gilt, wobei  $c \notin V$ . Da sowohl reguläre Sprachen als auch kontextfreie unter Konkatination und der Abbildung mit einem verallgemeinerten endlichen Automaten abgeschlossen sind, erweist sich  $\sigma_1(L)$  als regulär (kontextfrei), wenn  $L$  regulär (kontextfrei) ist.

Die Situation als Diagramm:



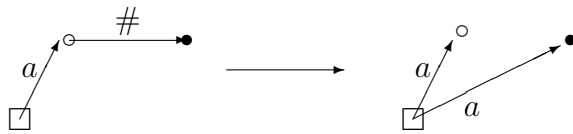
Es bleibt,  $gsm_R$  zu konstruieren. Sei dazu  $A_R$  ein endlicher Automat, der  $R$  erkennt, d.h.  $L(A_R) = R$ . Wegen der speziellen Form von Splicing-Regeln besitzt  $A_R$  folgende Gestalt:



Insbesondere gibt es in  $A_R$  keine Kreise, auf denen Sonderzeichen liegen, weil sonst Wörter mit mehr als drei Sonderzeichen erkannt würden.

In einem ersten Schritt wird  $A_R$  so umgebaut, dass alle Wörter der Form  $x_1 u_1 u_2 x_2 c y_1 u_3 u_4 y_2$  für  $x_1, x_2, y_1, y_2, y_z \in V^*$  und  $u_1\#u_2\$u_3\#u_4 \in R$  erkannt werden. Dazu werden zu  $A_R$  für alle  $a \in V$  Zustandsübergänge vom Anfangszustand, von den Quell- und Zielzuständen der  $\$$ -Kanten und von den Endzuständen jeweils zu sich selbst hinzugenommen. Außerdem werden die  $\$$ -Kanten durch  $c$ -Kanten ersetzt und die  $\#$ -Kanten gelöscht. Stattdessen wird für jede  $a$ -Kante, die die Quelle einer  $\#$ -Kante als Ziel hat, eine neue  $a$ -Kante von der alten Quelle

zum Ziel der  $\#$ -Kante als neuem Ziel gezogen:



Insbesondere erkennt dieser Automat ein Wort  $xcy \in L\{c\}L$  genau dann, wenn sich auf  $(x, y)$  eine Splicing-Regel anwenden lässt, d.h. wenn  $x = x_1u_1u_2x_2$  und  $y = y_1u_3u_4y_2$  für eine Regel  $u_1\#u_2\$u_3\#u_4 \in R$ . Da genau in diesem Fall  $z = x_1u_1u_4y_2 \in \sigma_1(L)$  ist, muss nur noch die Ausgabe so gewählt werden, dass bei Eingabe von  $x_1u_1u_2x_2cy_1u_3u_4y_2$  gerade  $x_1u_1u_4x_2$  ausgegeben wird. Das gelingt, wenn man die Ausgabe aller  $a$ -Kanten, die im  $u_1$ - und im  $u_4$ -Teil beginnen, identisch auf  $a$  setzt und aller anderen Zustandsübergänge auf  $\lambda$ .

Es muss beachtet werden, dass die bisherige Konstruktion voraussetzt, dass  $u_1 \neq \lambda \neq u_3$ . Für den Fall, dass  $u_1 = \lambda$  oder  $u_3 = \lambda$ , muss die Konstruktion modifiziert werden, was hier aber nicht explizit durchgeführt wird.