Theoretische Informatik
Kap 1: Formale Sprachen/Automatentheorie

Gliederung der Vorlesung

Wie weit tragen die Beschreibungsmittel für reguläre Sprachen?

- jede endliche Sprache ist regulär
- es gibt unendliche Sprachen, die regulär sind
- es gibt unendliche Sprachen, die nicht regulär sind

Beispiele:

$L = \{ 0^n 1^n \mid n > 1 \}$

$L' = \{ w \mid w \in \{ 0,1 \}^*, \#_0(w) = \#_1(w) \}$

Hinweis: $\#_0(w)$ und $\#_1(w)$ bezeichnen, wie oft eine 0 bzw. eine 1 in w vorkommen

Reguläre Sprachen

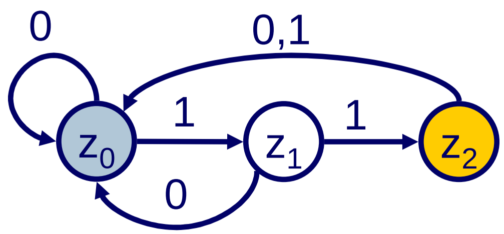
Was ist allen unendlichen regulären Sprachen gemein?

... um die Frage zu beantworten, muß man sich die zur Verfügung stehenden Beschreibungsmittel genauer ansehen

... die jeweils betrachteten Beschreibungsmittel haben Einfluß auf die Antworten und den Weg diese zu finden

... wählen deterministische endliche Automaten

Reguläre Sprachen



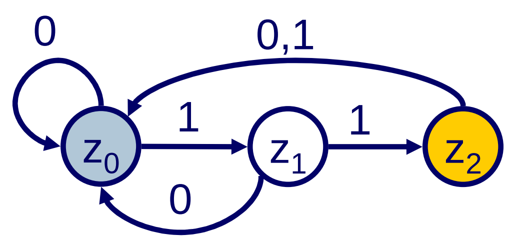
L(A) ist eine unendliche Sprachen, da ...

es sei s $\in$ L(A) mit $|s| \geq 3$

... dann muß der mit s markierte Weg durch den deterministischen endlichen Automaten A mindestens einen Knoten zweimal durchlaufen

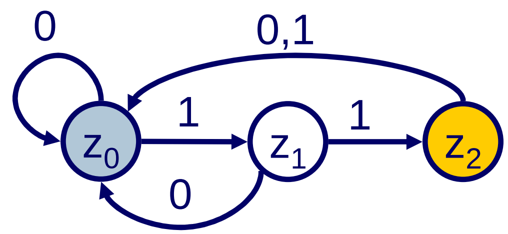
Reguläre Sprachen

markierte Wege	akzeptierte Wörter
$z_0z_0z_1z_2$	011
$z_0z_0z_0z_1z_2$	0011
$z_0z_1z_0z_1z_2$	1011
$z_0z_0z_0z_0z_1z_2$	00011
$z_0z_0z_1z_0z_1z_2$	01011
...	...



Reguläre Sprachen

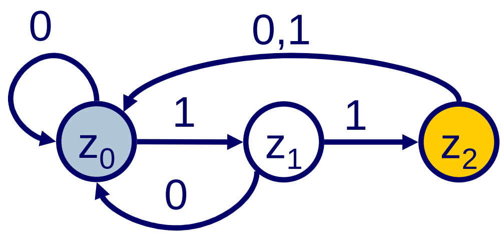
markierte Wege	**akzeptierte Wörter**
$z_0z_0z_1z_2$	**011**
$z_0z_0z_0z_1z_2$	**0011**
$z_0z_1z_0z_1z_2$	1011
$z_0z_0z_0z_0z_1z_2$	**00011**
$z_0z_0z_1z_0z_1z_2$	01011
...	...



L‘ = { 011, 0011, 000111, ... }
= { v^nw | v = 0, w = 11, n > 0 }

Reguläre Sprachen

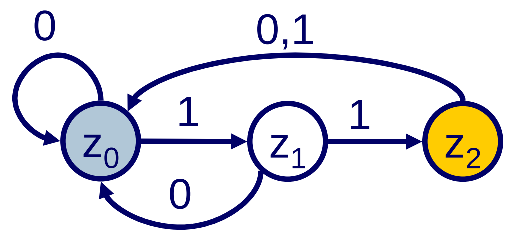
markierte Wege	**akzeptierte Wörter**
$z_0z_0z_1z_2$	011
$z_0z_0z_0z_1z_2$	0011
$z_0z_1z_0z_1z_2$	**1011**
$z_0z_0z_0z_0z_1z_2$	00011
$z_0z_0z_1z_0z_1z_2$	01011
...	...



$L' = \{ 1011, 101011, 10101011, \ldots \}$
$= \{ v^n w \mid v = 10, w = 11, n > 0 \}$

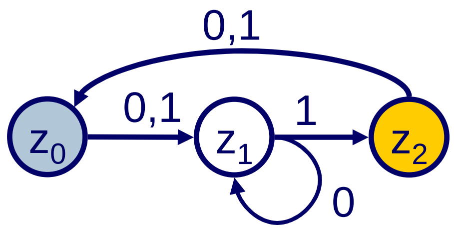
Reguläre Sprachen

markierte Wege	akzeptierte Wörter
$z_0z_0z_1z_2$	011
$z_0z_0z_0z_1z_2$	0011
$z_0z_1z_0z_1z_2$	1011
$z_0z_0z_0z_0z_1z_2$	00011
$z_0z_0z_1z_0z_1z_2$	**01011**
..	...



$$L' = \{ 01011, 001011, 0001011, \ldots \} = \{ v^n w \mid v = 0, w = 1011, n > 0 \}$$

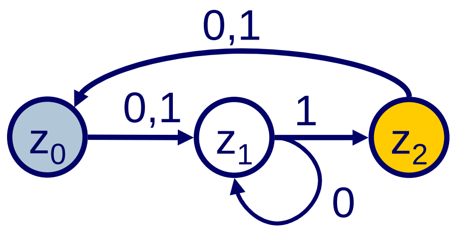
Reguläre Sprachen



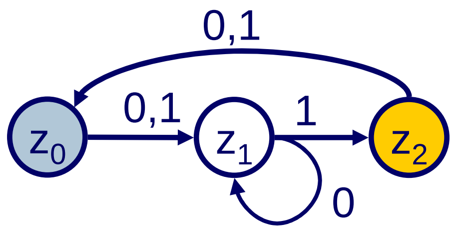
markierte Wege	akzeptierte Wörter
$z_0z_1z_1z_2$	001
$z_0z_1z_1z_2$	101
$z_0z_1z_1z_1z_2$	0001
$z_0z_1z_1z_1z_2$	1001
...	...
$z_0z_1z_2z_0z_1z_2$	11011
...	...

Reguläre Sprachen

markierte Wege	**akzeptierte Wörter**
$z_0z_1z_1z_2$	**001**
$z_0z_1z_1z_2$	101
$z_0z_1z_1z_1z_2$	**0001**
$z_0z_1z_1z_1z_2$	1001
...	...
$z_0z_1z_2z_0z_1z_2$	11011
...	...



$$L' = \{ 001, 0001, 00001, ... \}$$
$$= \{ uv^nw \mid u = 0, v = 0, w = 1, n > 0 \}$$

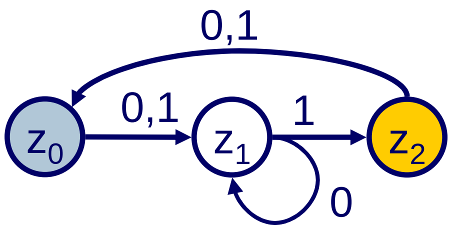


markierte Wege	akzeptierte Wörter
$z_0z_1z_1z_2$	001
$z_0z_1z_1z_2$	**101**
$z_0z_1z_1z_1z_2$	0001
$z_0z_1z_1z_1z_2$	**1001**
...	...
$z_0z_1z_2z_0z_1z_2$	11011
...	...

$$L' = \{ 101, 1001, 10001, ... \}$$
$$= \{ uv^nw \mid u = 1, v = 0, w = 1, n > 0 \}$$

Reguläre Sprachen

markierte Wege	**akzeptierte Wörter**
$z_0z_1z_1z_2$	001
$z_0z_1z_1z_2$	101
$z_0z_1z_1z_1z_2$	0001
$z_0z_1z_1z_1z_2$	1001
...	...
$z_0z_1z_2z_0z_1z_2$	**11011**
...	...



$$L' = \{ 11011, 11011011, 11011011011, ... \}$$
$$= \{ v^n w \mid v = 110, w = 11, n > 0 \}$$

Zwischenfazit

es sei A ein DFA mit n Zuständen

es sei L(A) eine unendliche Sprache

es sei s ein Wort aus L(A) mit $|s| \geq n$

dann gibt es Wörter u und w und ein Wort $v \neq \varepsilon$, so daß gilt:

- $u \cdot v \cdot w = s$
- $u \cdot v^k \cdot w \in L(A)$ für alle $k > 0$

... es gilt außerdem $|u| + |v| \leq n$ und $u \cdot w \in L(A)$

Reguläre Sprachen

allgemein

... was ist unendlichen regulären Sprachen L gemein

... es gibt ein DFA $A = [Z,\Sigma,\delta,z_0,E]$ mit $L(A) = L$

... es gibt ein $n > 0$ mit $card(Z) = n$

Pumping-Lemma für reguläre Sprachen

Es sei $L \subseteq \Sigma^*$ eine unendliche Sprache.

Wenn L regulär ist, so gibt es eine Zahl n,

so daß für alle Wörter $s \in L$ mit $|s| \geq n$, gilt:

es gibt Wörter u, v und w aus Σ^*, so daß gilt:

(i) $s = u \cdot v \cdot w$
(ii) $|v| \geq 1$
(iii) $|u| \cdot |v| \leq n$
(iv) $u \cdot v^k \cdot w \in L$ für alle $k \geq 0$

Was kann man mit dem Pumping-Lemma anfangen?

es sei $L \subseteq \Sigma^*$ eine unendliche Sprache

- falls L regulär ist,
  so gelten die Aussagen des Pumping-Lemma für L
- falls L nicht regulär ist,
  so gelten die Aussagen des Pumping-Lemma für L nicht

... das Pumping-Lemma kann demzufolge benutzt werden,
um nachzuweisen, daß L nicht regulär ist

Reguläre Sprachen

Was kann man mit dem Pumping-Lemma anfangen?

es sei $L \subseteq \Sigma^*$ eine unendliche Sprache

- falls L regulär ist, so gibt es <u>ein</u> n, so daß

  <u>für alle</u> $s \in L$ mit $|s| \geq n$ die im Pumping-Lemma formulierten Eigenschaften gelten

- falls L nicht regulär ist, so kann es <u>kein</u> n geben, so daß

  <u>für alle</u> $s \in L$ mit $|s| \geq n$ die im Pumping-Lemma formulierten Eigenschaften gelten

Reguläre Sprachen

es sei $L \subseteq \Sigma^*$ eine unendliche Sprache

- falls L nicht regulär ist, so kann es <u>kein</u> n geben, so daß

  <u>für alle</u> $s \in L$ mit $|s| \geq n$ die im Pumping-Lemma formulierten Eigenschaften gelten

- falls L nicht regulär ist, so muß es <u>zu jedem</u> n

  <u>ein</u> $s \in L$ mit $|s| \geq n$ geben, so daß für s die im Pumping-Lemma formulierten Eigenschaften <u>nicht</u> gelten

- falls L nicht regulär ist, so muß es <u>zu jedem</u> n

  <u>ein</u> $s \in L$ mit $|s| \geq n$ geben, so daß <u>für alle</u> u, v und w aus Σ^* <u>nicht</u> gleichzeitig die Bedingungen (i), (ii), (iii) und (iv) erfüllt sind

Reguläre Sprachen

Beispiel $L = \{ a^z b^z \mid z > 0 \}$

n	Kandidaten für s
0	ab, aabb, aaabbb, aaaabbbb, ...
1	ab, aabb, aaabbb, aaaabbbb, ...
2	aabb, aaabbb, aaaabbbb, ...
3	aabb, aaabbb, aaaabbbb, ...
4	aabb, aaabbb, aaaabbbb, ...
...	...

geeignete Auswahl treffen ... zu n wählen wir hier das Wort $s_n = a^n b^n$

Reguläre Sprachen

Beispiel $L = \{ a^z b^z \mid z > 0 \}$ (/* $n = 2$, $s_n = aabb$ */)

nachweisen, daß alle u, v und w aus Σ^* nicht gleichzeitig
(i), (ii), (iii) und (iv) erfüllen können

... aber es gibt unendlich viele u, v und w in Σ^*

... aber es gibt nur endlich viele u, v, w in Σ^* die gleichzeitig
(i), (ii) und (iii) erfüllen

(i) $s_n = u \cdot v \cdot w$
(ii) $|v| \geq 1$
(iii) $|u| \cdot |v| \leq n$
(iv) $u \cdot v^k \cdot w \in L$ für alle $k \geq 0$

Reguläre Sprachen

Beispiel $L = \{ a^z b^z \mid z > 0 \}$ (/* $n = 2$, $s_n = aabb$ */)

Idee: für die endlich vielen u, v und w aus Σ^*, die gleichzeitig (i), (ii) und (iii) erfüllen, zeigen, daß sie nicht (iv) erfüllen

Beobachtung: u, v und w aus Σ^* erfüllen (iv) nicht, falls gilt:

es gibt ein $k \geq 0$, so daß gilt: $u \cdot v^k \cdot w \notin L$

(i) $s_n = u \cdot v \cdot w$
(ii) $|v| \geq 1$
(iii) $|u| \cdot |v| \leq n$
(iv) $u \cdot v^k \cdot w \in L$ für alle $k \geq 0$

Reguläre Sprachen

Beispiel $L = \{ a^z b^z \mid z > 0 \}$ (/* $n = 2$, $s_n = aabb$ */)

wenn u, v, w aus Σ^* gleichzeitig (i), (ii) und (iii) erfüllen, liegt einer der folgenden Fälle vor

Fall 1: $v = a$

→ für $k = 0$, gilt: $u \cdot v^k \cdot w = abb \notin L$

Fall 2: $v = aa$

→ für $k = 0$, gilt: $u \cdot v^k \cdot w = bb \notin L$

(i) $s_n = u \cdot v \cdot w$
(ii) $|v| \geq 1$
(iii) $|u| \cdot |v| \leq n$
(iv) $u \cdot v^k \cdot w \in L$ für alle $k \geq 0$

Reguläre Sprachen

Beispiel $L = \{ a^z b^z \mid z > 0 \}$ (/* allgemein */)

für jedes $n > 0$ sei $s_n = a^n b^n$

es sei n beliebig, aber fest

wenn u, v und w aus Σ^* gleichzeitig (i), (ii) und (iii) erfüllen, so muß gelten:

- $v = a^r$ für ein r mit $0 < r \leq n$

also gilt für $k = 0$:

- $u \cdot v^k \cdot w = a^{n-r} b^n \notin L$

also ist L nicht regulär !!!

(i) $s_n = u \cdot v \cdot w$
(ii) $|v| \geq 1$
(iii) $|u| \cdot |v| \leq n$
(iv) $u \cdot v^k \cdot w \in L$ für alle $k \geq 0$

Reguläre Sprachen

es sei $L \subseteq \Sigma^*$ eine unendliche Sprache

um nachzuweisen, daß L nicht regulär ist, gehe wie folgt vor:

Schritt 1:

wähle zu jedem $n > 0$ ein $s_n \in L$ mit $|s_n| \geq n$ (/* geschickt */)

Schritt 2:

zeige für jedes $n > 0$ und alle u, v und w aus Σ^* mit:

$$s_n = u \cdot v \cdot w,\ |v| \geq 1 \text{ und } |u| \cdot |v| \leq n$$

daß für ein $k \geq 0$ gilt:

$$u \cdot v^k \cdot w \notin L$$

Beispiel $L = \{ ww \mid w \in \{ a,b \}^* \}$

für jedes $n > 0$ sei $s_n = a^n b a^n b$

es sei $n > 0$ beliebig, aber fest

wenn u, v und w aus Σ^* gleichzeitig (i), (ii) und (iii) erfüllen,
so muß gelten:

- $v = a^r$ für ein r mit $0 < r \leq n$

also gilt für $k = 0$:

- $u \cdot v^k \cdot w = a^{n-r} b a^n b \notin L$

also ist L nicht regulär !!!

Reguläre Sprachen

Begriffe:

- Chomsky-Grammatik, Ableitbarkeit
- reguläre Grammatik
- DFA, NFA

Sätze / Zusammenhänge:

- reguläre Grammatiken, DFA und NFA sind äquivalente Beschreibungsmittel für reguläre Sprachen
- Pumping-Lemma

Methoden / Techniken / Algorithmen:

- Übersetzung (/* reguläre Grammatiken $\leftrightarrow$ NFA $\leftrightarrow$ DFA */)
- Minimierungsalgorithmus
- Anwendung des Pumping-Lemma