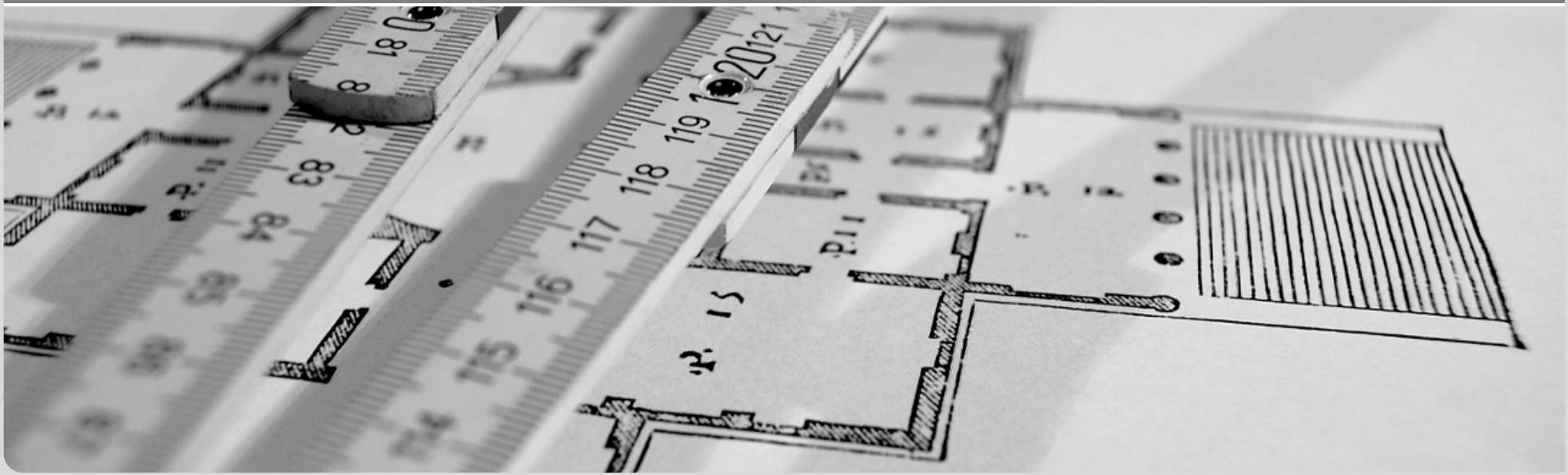
Grundbegriffe der Informatik
Tutorium 36

Termin 7 | 09.12.2016
Thassilo Helmold

KIT – Karlsruher Institut für Technologie



Inhalt

Relationen

Kontextfreie Grammatiken

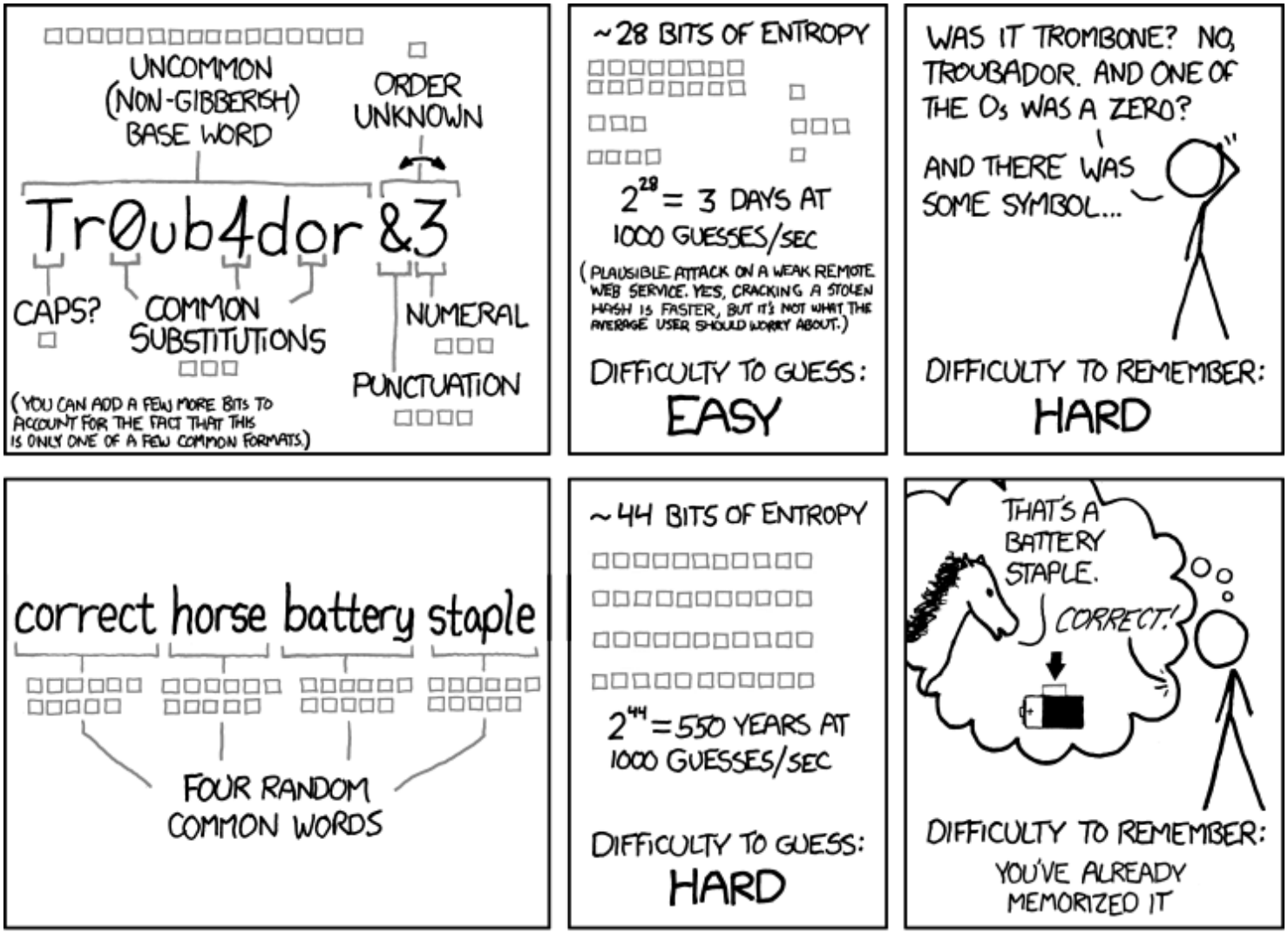


Abbildung: https://www.xkcd.com/936/

Zum letzten Übungsblatt

- Aufgabe 2.5: Die Abbildungen müssen wohldefiniert sein!
- Aufgabe 2.6: Def.-Bereich beachten (Alphabet oder Wort?)
- Aufgabe 2.6: Zielbereich: Surjektivität!
- Und noch einmal: Die Induktionsvorraussetzung latutet: Es gelte für EIN n ...

In the previous episode of GBI...

Rückblick: MIMA

- Ein idealisierter Prozessor
- Einfach zu verstehen, aufwändig zu programmieren
- Hardware-Details beachten: Keine negativen Konstanten mit LDC möglich!
- Programme sind oftmals mit „Bit-Magie“ einfacher und kürzer (aber auch schwerer zu verstehen)

Wahr oder Falsch?

- Alle Vögel haben die gleiche Farbe F (siehe Termin 6)
- Im IR der MIMA wird die Adresse des aktuellen Befehls gespeichert F
  Das geschieht im IAR. Im IR steht der Befehl selbst
- Die Befehlsholphase ist jede Ausführungsrunde identisch W
- Der Akku führt bei der MIMA Berechnungen aus F
  Das macht die ALU. Im Akku wird das letzte Ergebnis zwischengespeichert

Wahr oder Falsch?

- LDC -5 lädt -5 in den Akku F
  LDC funktioniert nicht mit negativen Konstanten!
- Mit der MIMA können wir Zufallszahlen erzeugen F
  Die MIMA arbeitet rein deterministisch und damit ohne jeden Zufall.

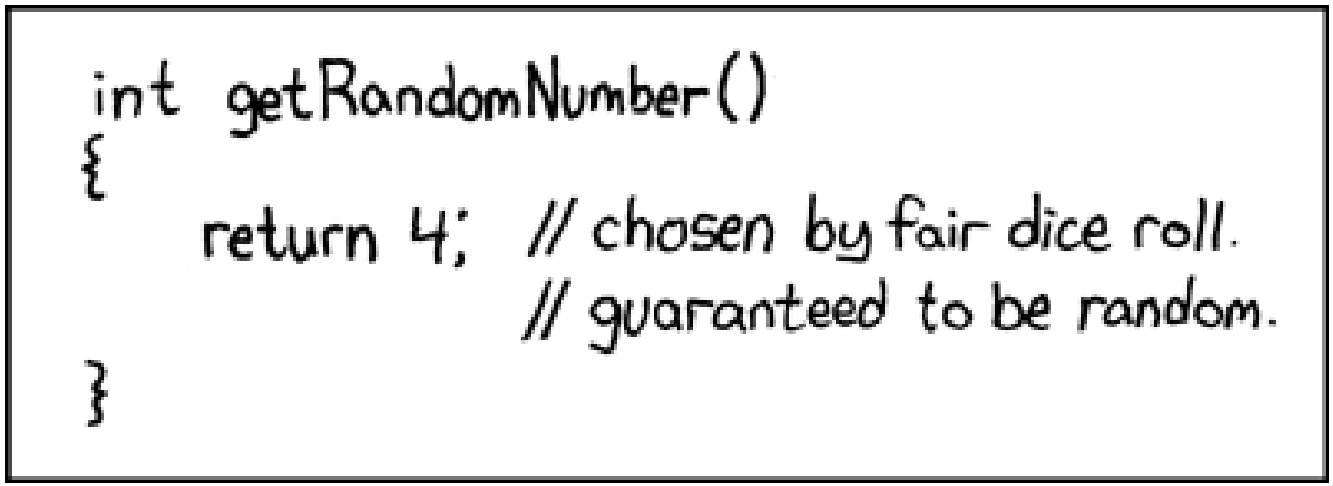


Abbildung: http://xkcd.com/

Der Philosoph am Scheideweg

Ein Weg führt in die Wüste, auf dem ein Wanderer marschiert. Nach einem eintägigen Fußmarsch gelangt er an eine Weggabelung, an der eine kleine Hütte steht.

Neben der Eingangstüre ist ein Plakat mit folgendem Text befestigt: Einer der beiden weiterführenden Wege führt zu mehreren guten Oasen; nur auf diesem Weg kannst du dein Ziel lebend erreichen.

Der andere Weg führt in die unendliche Wüste; er hat noch jedem den Tod gebracht, der ihn wählte.

Der Philosoph am Scheideweg

In dieser Hütte haust ein Philosoph, der die Ziele der beiden Wege sicher kennt; **streng abwechselnd sagt er an einem Tag die Wahrheit und am darauf folgenden Tag lügt er**.

Keiner außer ihm weiß aber, ob der heute die Wahrheit sagt oder aber heute lügt.

Du darfst dem Philosophen nur eine einzige Frage stellen, mit der du den Weg zu den Oasen erfahren willst.

Was fragst du ihn?

Der Philosoph am Scheideweg

Was fragst du ihn?

Was würdest du sagen, wenn ich dich morgen fragen würde, ob der linke Weg der richtige ist?

Und wenn der Philosoph jeden Morgen würfelt, ob er die Wahrheit sagt oder lügt?

Der Philosoph am Scheideweg

Was fragst du ihn?

Was würdest du sagen, wenn ich dich morgen fragen würde, ob der linke Weg der richtige ist?

Und wenn der Philosoph jeden Morgen würfelt, ob er die Wahrheit sagt oder lügt?
Was würdest du sagen, wenn ich dich **heute** fragen würde, ob der linke Weg der richtige ist?

Relationen

Kontextfreie Grammatiken

Eigenschaften

Definition

Sei $R \subseteq A \times A$ eine (binäre) Relation auf der Menge A. Wir nennen R

- **reflexiv** falls gilt

$$\forall x \in A : (x, x) \in R$$

- **symmetrisch** falls gilt

$$\forall x, y \in A : (x, y) \in R \implies (y, x) \in R$$

- **transitiv** falls gilt

$$\forall x, y, z \in A : (x, y) \in R \text{ und } (y, z) \in R \implies (x, z) \in R$$

Beispiele

- Die Relation $=$ ist reflexiv, symmetrisch und transitiv. Man nennt so etwas auch Äquivalenzrelation
- Die Relation $<$ ist nicht reflexiv und nicht symmetrisch, aber transitiv
- Die Relation $\leqslant$ ist reflexiv, nicht symmetrisch, aber transitiv

Produkt

Definition

Das **Produkt** von zwei Relationen $R \subseteq M \times N, S \subseteq N \times L$ definieren wir als

$$S \circ R = \{(x, z) \in M \times L \mid \exists y \in N \ : \ (x, y) \in R \text{ und } (y, z) \in S\}$$

Definition

Die **Potenz** einer Relation $R \subseteq M \times M$ definieren wir als

$$R^0 = I_M = \{(x, x) \mid x \in M\}$$
$$R^{i+1} = R^i \circ R$$

Beobachtung

Wenn f und g Funktionen sind (also linkstotale, rechtseindeutige Relationen), entspricht $f \circ g$ der Hintereinanderausführung von f nach g.

Reflexiv-transitive Hülle

Definition

Die **reflexiv-transitive Hülle** einer Relation R ist

$$R^* = \bigcup_{i=0}^{\infty} R^i$$

Satz

R^* ist die kleinste Relation, die R umfasst und reflexiv und transitiv ist.

Beispiel

Sei $A = \{a, b, c, d, e\}$ und R = $\{(a, b), (b, c), (c, e)\} \subseteq A \times A$
$R^* = \{(a, a), (b, b), (c, c), (d, d), (e, e),$
$(a, b), (b, c), (c, e),$
$(a, c), (b, e), (a, e)\}$

Noch offen: Klammerausdrücke

A long, long time ago, in a land far away:
Formale Sprachen angeben durch Mengen, Konkatenation und Kleenschem Abschluss...
Was ist mit der Sprache aller gültigen Klammerausdrücke? Können wir diese auch auf diese Weise angeben?

Jetzt wissen wir: Nein, das geht nicht! (Siehe VL)

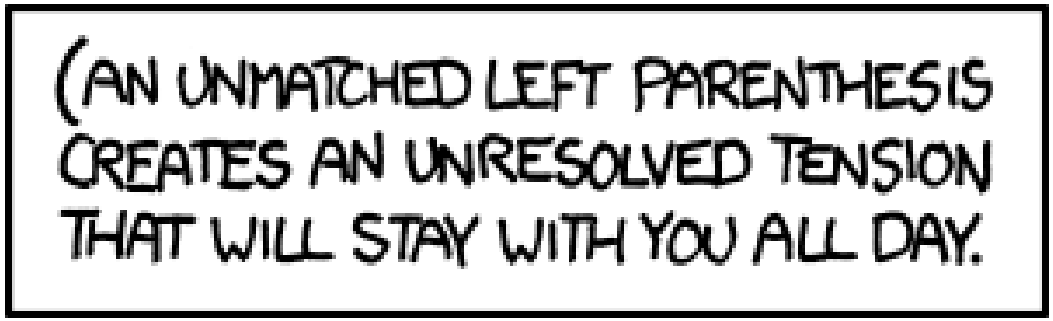


Abbildung: `https://xkcd.com/859/`

Relationen

Kontextfreie Grammatiken

Kontextfreie Grammatiken

Definition

Eine **kontextfreie Grammatik** ist ein 4-Tupel $G = (N, T, S, P)$ mit

- N Alphabet von Nichtterminalsymbolen
- T Alphabet von Terminalsymbolen ($N \cap T = \emptyset$)
- S Startsymbol ($S \in N$)
- P Produktionsmenge ($P \subseteq N \times (N \cup T)^*$)

Beispiel

Sei A das deutsche Alphabet (mit Klein-/Großbuchstaben)
$G_{MI} = (\{S, M, I, N\}, A \cup \mathbb{N}_+, S, P)$

$$
\begin{aligned}
P = \{S &\to \mathrm{M\_I\_N}, \\
M &\to \mathit{Monkey}, \\
I &\to \mathit{Island}, \\
N &\to 1 \mid 2 \mid 3\}
\end{aligned}
$$

Kontextfreie Grammatiken

Produktionen

Menge von gültigen Ersetzungen. Ersetzt wird immer **genau ein Nichtterminal** (kontextfrei) mit einem Wort von Zeichen aus $(N \cup T)$.
Vereinfachende Schreibweise:

$$S \rightarrow a \mid b$$

(S kann durch a oder b ersetzt werden)

Ableitung

Definition

Für ein $u \in V^*$ mit $u = w_1 \cdot X \cdot w_2$ $(w_1, w_2 \in V^*, X \in N)$ nennen wir ein Wort $v = w_1 \cdot w \cdot w_2 \in V^*$ **ableitbar**, wenn eine Produktion $X \to w$ existiert.
Wir schreiben

$$u \Rightarrow v$$

Beispiel

Für G_{MI} gilt: $S \Rightarrow$ M␣I␣N
M␣I␣N $\Rightarrow$ Monkey␣I␣N
Es gibt kein Wort, das aus Monkey Island 1 abgeleitet werden kann.

Ableitung

Definition

$$u \Rightarrow^0 v \text{ genau dann, wenn } u = v$$
$$u \Rightarrow^{i+1} v \text{ genau dann, wenn für ein } w \in V^* : u \Rightarrow w \Rightarrow^i v$$
$$u \Rightarrow^* v \text{ genau dann, wenn für ein } i \in \mathbb{N}_0 : u \Rightarrow^i v$$

Beobachtung

Die Definitionen stimmen mit den Potenzen der Relation $\Rightarrow$ überein.
$\Rightarrow^*$ ist die reflexiv-transitive Hülle von $\Rightarrow$.

Beispiel

Für G_{MI} gilt: $S \Rightarrow^2$ Monkey Island N $\Rightarrow$ Monkey Island 3
$S \Rightarrow^*$ Monkey Island 2

Erzeugte Sprache

Definition

Sei G eine kontextfreie Grammatik. Wir bezeichnen die Sprache

$$L(G) = \{w \in T^* \mid S \Rightarrow^* w\} \subseteq T^*$$

als die von der Grammatik G **erzeugte Sprache**.

Das sind also alle Wörter aus Terminalsymbolen, die vom Startsymbol aus ableitbar sind.
Achtung: Die erzeugte Sprache kann auch leer sein.
Beispiel: $L\left((\{X\}, \{a, b\}, X, \{X \to X\})\right) = \{\}$

Erzeugte Sprache

Beispiel

$L(G_{MI}) = \{\text{Monkey Island 1, Monkey Island 2, Monkey Island 3}\}$
$\text{M I N} \notin L(G_{MI})$

Definition

Eine Sprache L, für die eine kontextfreie Grammatik G mit $L(G) = L$ existiert, heißt **kontextfrei**.

Viele „natürlich vorkommende“ Sprachen sind kontextfrei.

Beispiel

$$G = (\{X\}, \{a, b\}, X, \{X \to aXb \mid \varepsilon\})$$

- Gilt $X \Rightarrow aXb$, $X \Rightarrow aaXbb$, $XX \Rightarrow aXbaXb$?
- Welche Wörter lassen sich aus aaXbb ableiten? Und aus XX?
- Gib $L(G)$ an!

Musikgrammatik

Wir betrachten die Grammatik

$$G = (\{X\}, \{A, B, C, D\}, X, \{X \to \varepsilon \mid AX \mid BX \mid CX \mid DX\})$$

Wie kann man A ableiten? Und ABC? Und

$$X \Rightarrow AX \Rightarrow ABX \Rightarrow ABBX \Rightarrow ABBAX \Rightarrow ABBA$$

Musikgrammatik

$$G = (\{X\}, \{A, B, C, D\}, X, \{X \to \varepsilon \mid AX \mid BX \mid CX \mid DX\})$$

Welche Wörter kann man nicht ableiten?

Musikgrammatik

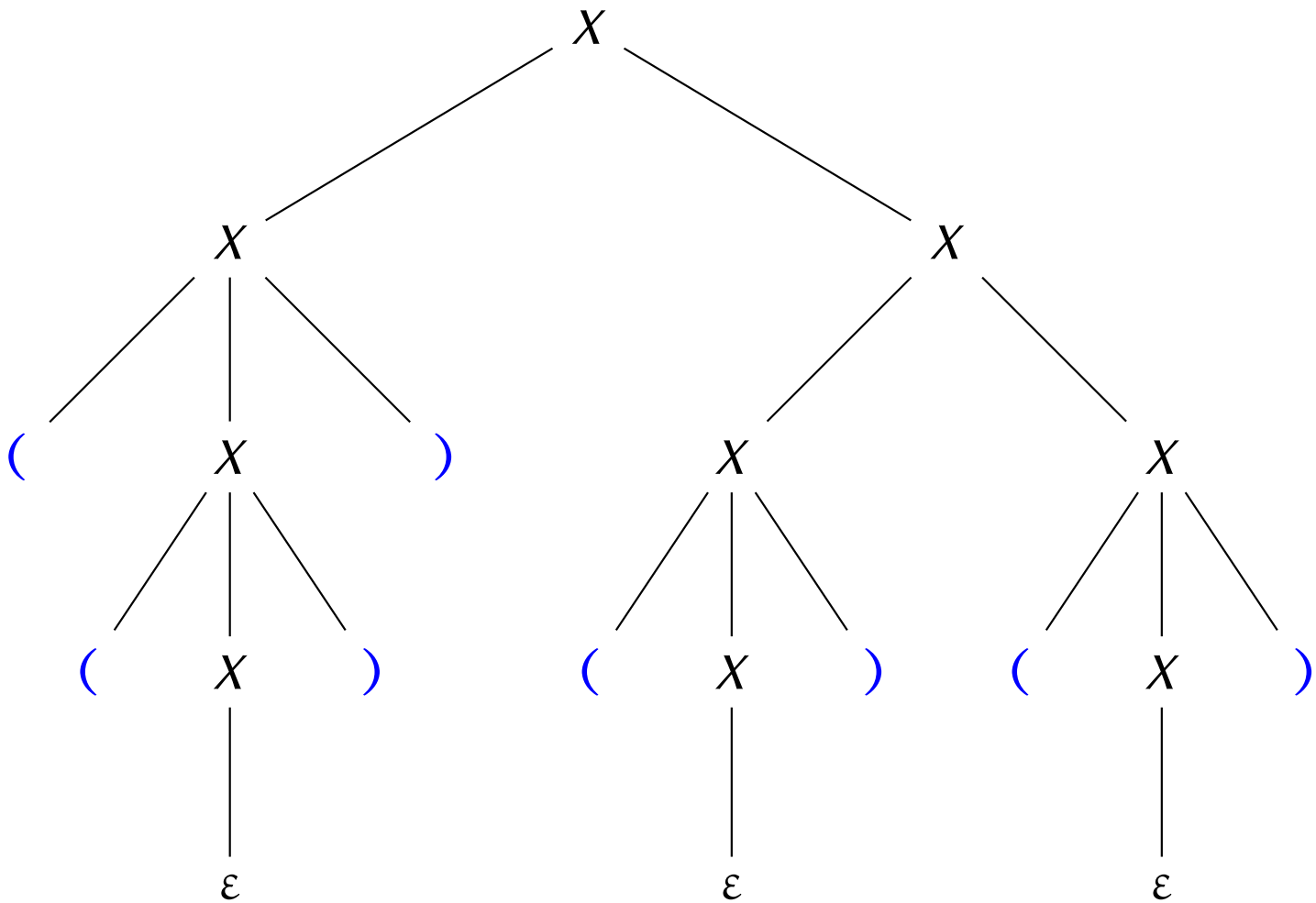
$$G = (\{X\}, \{A, B, C, D\}, X, \{X \to \varepsilon \mid AX \mid BX \mid CX \mid DX\})$$

Welche Wörter über $\{A, B, C, D\}$kann man nicht ableiten?
Keine! Die erzeugte Sprache ist

$$L(G) = \{A, B, C, D\}^*$$

Ableitungsbäume

$$G = (\{X\}, \{(,)\}, X, \{X \to XX \mid (X) \mid \varepsilon\})$$



Klammerausdrücke

Gegeben sei die Grammatik

$$G = (\{X\}, \{(,)\}, X, \{X \to XX \mid (X) \mid \varepsilon\})$$

- Wie leitet man $((((()))))$ ab?
- Wie leitet man $()(())()()$ ab?
- Kann man $(()($ ableiten? Nein!

Klammerausdrücke

Gegeben sei die Grammatik

$$G = (\{X\}, \{(,)\}, X, \{X \to XX \mid (X) \mid \varepsilon\}$$

Was ist $L(G)$? Was kann man also aus X ableiten?
Alle „*wohlgeformten Klammerausdrücke*“

Was bedeutet wohlgeformt in diesem Kontext?

$$\forall w \in L(G) : N_{(}(w) = N_{)}(w)$$

Reicht das? Notwendig aber nicht hinreichend!

Klammerausdrücke

Wir dürfen eine Klammer erst schließen, *nachdem* wir sie geöffnet haben.
Also: Anzahl der schließenden Klammern darf nie größer als Anzahl der öffnenden Klammern sein!

Für jedes Präfix v von einem Wort $w \in L(G)$ gilt

$$N_{(}(v) \geqslant N_{)}(v)$$

Achtung: Grammatiken sind nicht eindeutig! Wir können zur gleichen Sprache mehrere verschiedene erzeugende Grammatiken finden.
Alternative Grammatik für wohlgeformte Klammerausdrücke:

$$G = (\{X\}, \{(,)\}, X, \{X \to (X)X \mid \varepsilon\})$$

Und jetzt ihr...

Gebt jeweils eine Grammatik über dem Alphabet $T = \{a, b\}$ an, die folgende Sprache erzeugt:

- Alle Wörter, in denen irgendwo das Teilwort *baa* vorkommt.

- Die Menge aller Wörter $w \in T^*$ mit der Eigenschaft, dass für alle Präfixe v von w gilt: $|N_a(v) - N_b(v)| \leqslant 1$.
  Tipp: Was für eine Struktur haben Wörter der Länge 2, 4, ...?

- Alle Wörter, in denen *ab* als Teilwort vorkommt oder kein *a* enthalten ist.

Und jetzt ihr...

Gebt jeweils eine Grammatik über dem Alphabet $T = \{a, b\}$ an, die folgende Sprache erzeugt:

- Alle Wörter, in denen irgendwo das Teilwort baa vorkommt.
  $(\{X, Y\}, T, X, P)$ mit $P = \{X \to YbaaY, Y \to aY|bY|\varepsilon\}$
- Die Menge aller Wörter $w \in T^*$ mit der Eigenschaft, dass für alle Präfixe v von w gilt: $|N_a(v) - N_b(v)| \leqslant 1$.
  Tipp: Was für eine Struktur haben Wörter der Länge 2, 4, ...?
  $(\{X, Y\}, T, X, P)$ mit $P = \{X \to abX|baX|a|b|\varepsilon\}$
- Alle Wörter, in denen ab als Teilwort vorkommt oder kein a enthalten ist.
  $G = (\{X, Y\}, \{a, b\}, X, P)$ mit $P = \{X \to bX \mid YabY \mid \varepsilon, Y \to aY \mid bY \mid \varepsilon\}$

Aufgabe (WS 2008)

- Geben Sie eine kontextfreie Grammatik

$$G = (N, \{a, b\}, S, P)$$

  an, für die $L(G)$ die Menge aller Palindrome über dem Alphabet $\{a, b\}$ ist.
- Geben Sie eine Ableitung der Wörter **baaab** und **abaaaba** aus dem Startsymbol Ihrer Grammatik an.
- Beweisen Sie, dass Ihre Grammatik jedes Palindrom über dem Alphabet $\{a, b\}$ erzeugt.
  Tipp: Induktion: Wenn n und $n+1$ gelten, dann gilt auch $n+2$

Lösung

Die Grammatik

$$G = (\{S\}, \{a, b\}, S, P = \{S \to aSa \mid bSb \mid a \mid b \mid \varepsilon\})$$

erzeugt gerade die Menge der Palindrome. Die Ableitungen der Wörter mit dieser Grammatik sind

$$S \Rightarrow bSb \Rightarrow baSab \Rightarrow baaab$$

$$S \Rightarrow aSa \Rightarrow abSba \Rightarrow abaSaba \Rightarrow abaaaba$$

Lösung

Sei w ein Palindrom über $\{a, b\}$. Wir zeigen durch Induktion über $n = |w|$, dass alle Palindrome aus S abgeleitet werden können.

Induktionsanfang

Für $n = 0$ ist das leere Wort ε in einem Schritt aus S ableitbar.
Für $n = 1$: Die einzigen Wörter aus $\{a, b\}^*$ der Länge 1 sind a und b. Auch diese sind offensichtlich aus S ableitbar.

Induktionsvoraussetzung

Für ein festes, aber beliebiges $n \in \mathbb{N}_0$ gilt, dass alle Palindrome der Länge n und alle Palindrome der Länge $n+1$ aus S abgeleitet werden können.

Lösung

Induktionsschritt

Sei w ein Palindrom der Länge $n+2$. Das erste (und damit auch das letzte) Zeichen sei oBdA ein a. Dann gibt es ein $w' \in \{a, b\}^*$, so dass $w = aw'a$ ist. Da w ein Palindrom ist, muss auch w' ein Palindrom sein. Weiterhin gilt $|w'| = n$. Nach IV gibt es somit eine Ableitung $S \Rightarrow^* w'$. Somit gibt es die Ableitung

$$S \Rightarrow aSa \overset{IV}{\Rightarrow}^* aw'a = w$$

und $w \in L(G)$ folgt. Entsprechendes gilt, wenn das erste Zeichen von w ein b ist.
Mit der IV haben wir also gezeigt, dass auch Palindrome der Länge $n+1$ und $n+2$ aus S ableitbar sind.

Gibt es noch mehr?

„Die meisten“ Sprachen in der Informatik sind kontextfrei.

Was ist mit der Sprache $L_{vv} = \{vcv \mid v \in \{a, b\}^*\}$
In der Vorlesung: Es gibt keine kontextfreie Grammatik, die L_{vv} erzeugt.
Können wir die Sprache trotzdem irgendwie „verarbeiten“?

Soon

Was ihr nun wissen solltet

- Mehr Eigenschaften von Relationen
- Was eine kontextfreie Grammatik ist
- Wie man Sprachen aus Grammatiken ableiten kann

Was nächstes Mal kommt

- Nachts sind alle Katzen grau - Prädikatenlogik
- Algorithmen: Kochrezepte der Informatik

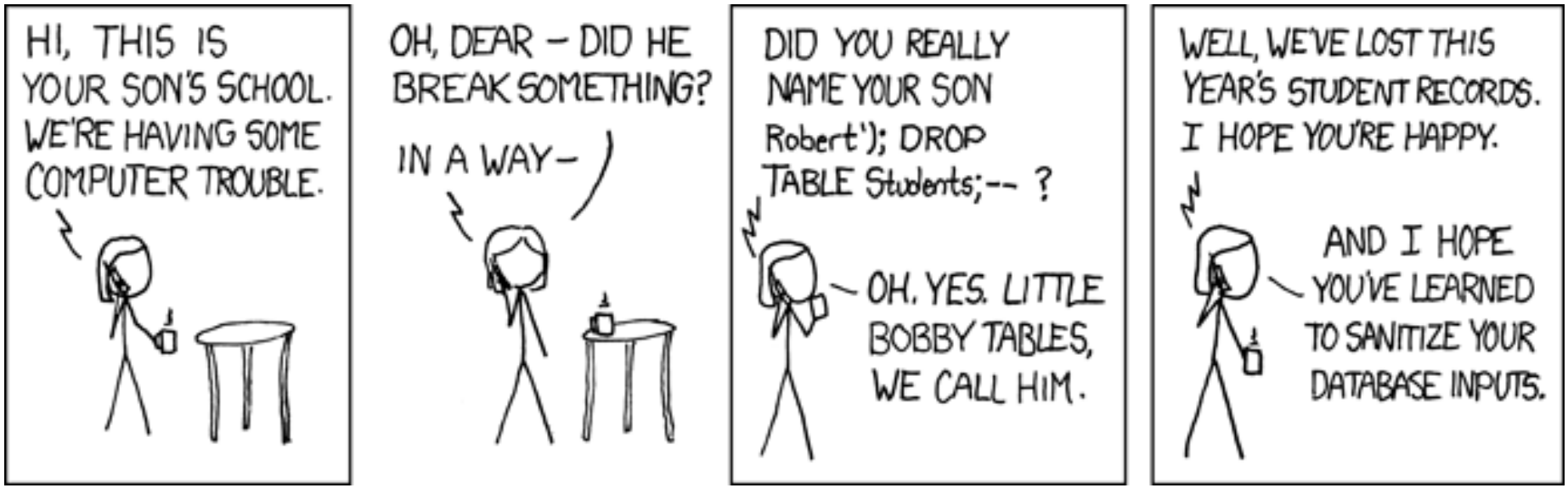


Abbildung: https://www.xkcd.com/327/

Danksagung

Dieser Foliensatz basiert in Teilen auf Folien von:

Philipp Basler
Nils Braun
Dominik Doerner
Ou Yue