Predicate:

a mathematical predicate can be True or False.
predicates are functions with boolean return values:
P, Q(n), R(x,y,z)

Logical Operators:

AND: $P \land Q$ || OR: $P \lor Q$ || NOT: $\neg P$
Implication: $P \implies Q = \neg P \lor Q$

Distributive Rule

$P \land (Q \lor R) = (P \land Q) \lor (P \land R)$
$P \lor (Q \land R) = (P \lor Q) \land (P \lor R)$

De Morgans Law

$\neg(P \land Q) = \neg P \lor \neg Q$
$\neg(P \lor Q) = \neg Q \implies \neg P$
$P \implies Q = \neg Q \implies \neg P$

Implication

$P \implies Q = True$ & $\neg P \implies Q = True$
$P \implies \neg Q = False$
$\neg P \implies \neg Q = True$

Quantors

OR: $\bigvee_{k=0}^{n} P_k$
P true for any $k \in 0..n$
All: $\forall k \in 0..n = P_k$
for all k P = True

AND: $\bigwedge_{k=0}^{n} P_k$
P true for all $k \in n$
Exists: $\exists k \in 0..n = P_k$
a k exists where P = True

Normalforms

disjunctive
$(x1 \land x2) \lor (\overline{x1} \land x2) \lor (x1 \land \overline{x2})$

conjunctive
$(x1 \lor x2) \land (\overline{x1} \lor x2) \land (x1 \lor \overline{x2})$

These are useful for true and false tables
This one would result to true if x1 or x2 is true.

Quantities:

$\emptyset = \{\}$ $[n] = \{0..n\}$ {a .. z}

Union: $A \cup B = \{x | x \in A \lor x \in B\}$
Intersection: $A \cap B = \{x | x \in A \land x \in B\}$
Complement: $\overline{A} = \{x | x \notin A\}$
Difference: $A \setminus B = \{x \in A | x \notin B\}$

Pairs $A \times B = \{(a,b) | a \in A \land b \in B\}$

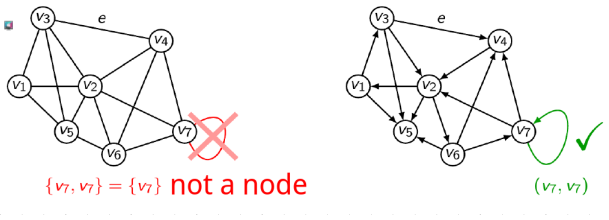
n-Tuples $\bigtimes_{k=0}^{n} A_i = \{(a_o, a_i, .., a_n) | a_i \in A_i\}$

Undirected Graph

doesn't have directions, and therefore can't have edges to itself

Directed Graph

This does have directions, therefore an edge to itself is valid!



Vertices V = $\{v_1, v_2, ..., v_n\}$ Edge e = $\{v_3, v_4\}$
EdgeCount E = $\{e | e Edge\}$

Proofs:

constructive Proof (proof by reforming)

Consider $ax^2 + bx + c = 0$, we can proof this to have 2 solutions by reforming.

$$
\begin{aligned}
ax^2 + bx + c &= 0 \\
x^2 + 2\frac{b}{2a}x + \frac{b^2}{4a^2} - \frac{b^2}{4a^2} + \frac{c}{a} &= 0 \\
x^2 + 2\frac{b}{2a}x + \frac{b^2}{4a^2} &= \frac{b^2 - 4ac}{4a^2} \\
\left(x + \frac{b}{2a}\right)^2 &= \frac{b^2 - 4ac}{4a^2} \\
x + \frac{b}{2a} &= -\frac{b}{2a} \pm \sqrt{\frac{b^2 - 4ac}{4a^2}}
\end{aligned}
$$

If $b^2 - 4ac > 0$ then we have 2 solutions!

Proof by contradiction

Take -2, it isn't a natural number. We can prove this by claiming the opposite.
If -2 is a natural number, then it has all the attributes of a natural number.
For example, it should be possible to take the square root of -2.
$\sqrt{-2} = NaN$
As you can see -2 does not have this attribute and is therefore
NOT a natural number!

Proof by Induction

This is particularly useful if you want to check an attribute
for a range of numbers such as n or n+1

Base claim:
$P(n) = \sum_{k=1}^{n} = \frac{n(n+1)}{2}$

Anker: check for n=1
$P(1) = \frac{1(1+1)}{2} = 1$

Hypothesis: it also works for n+1

$$
\begin{aligned}
P(n+1): \quad \sum_{k=1}^{n+1} k &= \left(\sum_{k=1}^{n} k\right) + n + 1 \\
&= \frac{(n+1)(n+2)}{2} \\
&= \frac{(n+1)(n+1+1)}{2} \quad \checkmark
\end{aligned}
$$

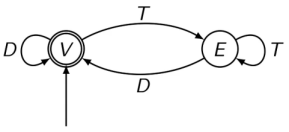
Alphabet and Word

Σ = Alphabet: Nonempty Quantity of characters
$\Sigma^n = \Sigma \times ... \times \Sigma$ = String
$w \in \Sigma^n$ An element in that string is a Word with length n.
$\varepsilon \in \Sigma^0$ The empty word, don't forget the empty word!
Quantity of all words:
$\Sigma^* = \{\varepsilon\} \cup \Sigma \cup \Sigma^2 \cup \Sigma^3 \cup ... = \bigcup_{k=0}^{\infty} \Sigma^k$

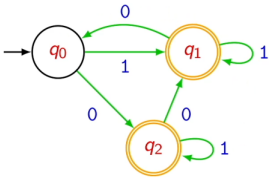
Language $L \subset \Sigma^*$ = Language

$L = \emptyset \subset \Sigma^*$ = Empty Language
$L = \Sigma^* \to \Sigma = \{0, 1\}$ all binary strings.
A language is regular if a DFA can be formed out of it.

Deterministic Finite Automaton (DFA/DEA)



A very simple machine that accepts a variety of inputs.
Only requirement is that a D follows after T.
This means all the following inputs are valid:
_ (empty word!), D, DD, TD, TTTTTTTD, DDDDDTD, DDDDD, ...



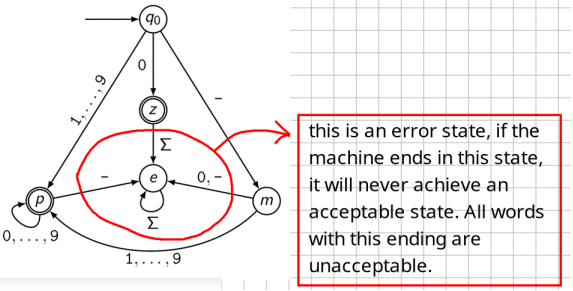
Machine A: {Q,Σ,δ,q0,F}
State = Q -> $\{q_1, q_2, ..., q_n\}$
Alphabet = Σ
Transitioning-Function = $\delta : Q \times \Sigma \to Q$
Starting State = $q_0 = L(\varepsilon) = L$
Acceptable Endstates = $F \subset Q$

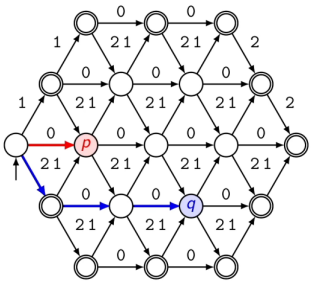
	0	1	
all states Q → q_0	q_2	q_1	← Σ possible inputs, here 0 or 1
q_1 /F	q_0	q_1	← δ current position. -> change to q1
all acceptable states. F → q_2 /F	q_1	q_2	

Language of DFA A:

$L(A) = \{w \in \Sigma^* | A \text{ accepts } w\} = \{w \in \Sigma^* | \delta(q_0, w) \in F\}$
The language of a DFA is simply all accepted words!

Error States in DFA





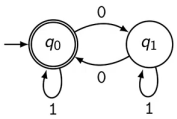
from q, many paths lead to F
This means 11, or 0 would be the "same"
$L(q) = \{0, 10, 11, 12, ...\}$
The same would obviously apply to P

Myhill-Nerode

Adding a word to a word, to make it compatible with a language

$L(w) = \{w' | ww' \in L\}$ including: $L(\varepsilon)$!

w	$L(w)$	Q
ε	$L(\varepsilon) = L$	q_0
0	$L(0) = \{w \in \Sigma^* \mid \lvert w\rvert_0$ uneven $\}$	q_1
1	$L(1) = \{w \in \Sigma^* \mid \lvert w\rvert_0$ even $\} = L$	q_0
⋮	⋮	⋮



even and uneven amounts of 0s

mod 2 zero's, 1's don't matter

Detecting Nonregular Languages with Myhill

The examples before always had a specific amount of words/characters that one had to add, in order to accept the word.

However, there are languages that would need infinite states in order to find the entire language of a DFA

A good example for this is the language 1^n0^n

w	$L(w)$	Q
ε	$\{0^n1^n \mid n \geq 0\}$	q_0
0	$\{0^n1^{n+1} \mid n \geq 0\}$	q_1
00	$\{0^n1^{n+2} \mid n \geq 0\}$	q_2
000	$\{0^n1^{n+3} \mid n \geq 0\}$	q_3
0^k	$\{0^n1^{n+k} \mid n \geq 0\}$	q_k
	⋮	
01	$\{\varepsilon\}$	
001	$\{1\}$	
0001	$\{11\}$	
	⋮	
1	$\emptyset$	e
10	$\emptyset$	e
	⋮	

for every 0 that we add, we need a 1

this means that for n+k 0's we need k 1's

as $\lim_{k\to\infty}$ we need ∞ states!

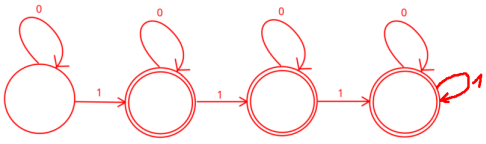
not possible with a Deterministic automaton!

also note: we have clear error states

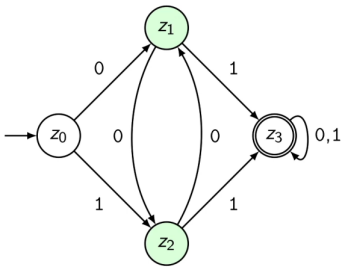
anything starting with 1 is an error.

Differentiation of States

To get a minimal DFA, we eliminate all states that are superfluous.



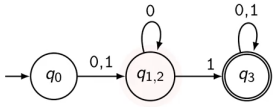
Here the 3 acceptable states can be put together, they are the same!



	z_0	z_1	z_2	z_3
z_0	≡	×	×	×
z_1	×	≡	≡	×
z_2	×	≡	≡	×
z_3	×	×	×	≡

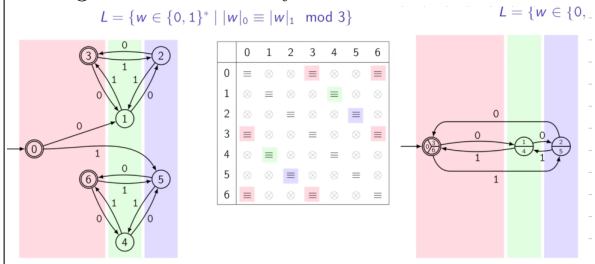
z1 and z2 are the same!

they can't be differntiated



New minimal Automaton!

This algorithm makes it easy to see whether or not states are the same!



Beauty of a language / Pumping Lemma

a language L can be pumped if the following is valid:

$\exists N > 0$ where $w \in L \wedge |w| \geq N$

If this word can be divided into 3 parts x,y,z while:

$|xy| \leq N \wedge |y| > 0 \wedge |x| \geq 0 \wedge |z| \geq 0 \wedge xy^kz \in L \rightarrow \forall k \in \mathbb{N}$

Find a number N that is bigger than 0 but less than the length of xy

while the length of y is greater than 0 and xy^kz is still part of the language

Note: The power of xy^kz is NOT a power, but an indicator how many y's!!!

Any language that can be pumped is a regular language, and is therefore "schön"

Consider the following 2 examples:

$L(s1) = \{0, 1 | \text{ending with } 1\}$

z = 1, xy = any combination of 1 and 0

try: x=0 y=1010111 z=1,True -> $N \geq 7$

try: x=0 y=1 z=1,True -> $N \geq 1$

This can be done with any y, x

It will always satisfy all requierements of the pumping lemma.

this language is regular.

$L(s1) = \{0^n, 1^n | n \geq 0\}$

if z=1, more 1's than 0's, FALSE

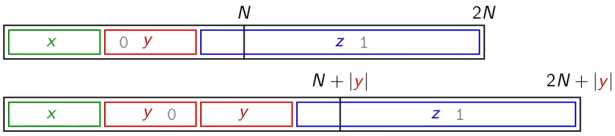
if x - 0, more 0's than 1's, FALSE

This means we need x=0, z=0

If we do that, then y is empty

No matter what we do, it is FALSE

0^n1^n is therefore not regular



Pumping Lemma usage guide !!FOLLOW THIS!!

1. Claim that L is regular
2. According to pumping lemma $\exists N$

Don't make claims about the size of N!

3. Choose a word $w \in L || w| \geq N$

Definition with N has to be written!

4. Division into parts according to Pumping Lemma

$w = xyz, |xy| \leq N, |y| > 0$,etc

5. Check if word is in language

min 1 word not in language: $xy^kz \notin L | k \in \mathbb{N}$

Explain why this word is not in the language

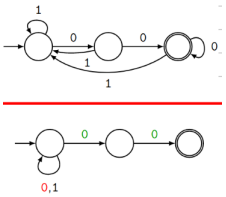
6. Contradiction, aka explain that this language is not regular.

Non-Deterministic Finite Automaton (NFA / NEA)

Before every step was clear, there was no other determinism other than the input and the current state. A non-deterministic automaton

Can have other things, like only accept the last 2 0's

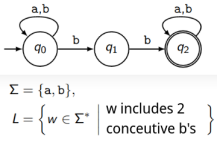
This is illustrated as both a DFA and an NFA:



Both have the same goal, only the last 2 zero's lead to the acceptance state.

However one is obviously easier, while not giving clear info on in what state it is, it is hence non-deterministic.

Formal Definition:



Machine A: {Q,Σ,δ,q_0,F}

State = Q -> $\{q_1, q_2, ..., q_n\}$

Alphabet = Σ

Transitioning-Function = $\delta : Q \times \Sigma \rightarrow \mathbf{P(Q)}$

Starting State = $q_0 = L(\varepsilon) = L$

Acceptable Endstates = $F \subset Q$

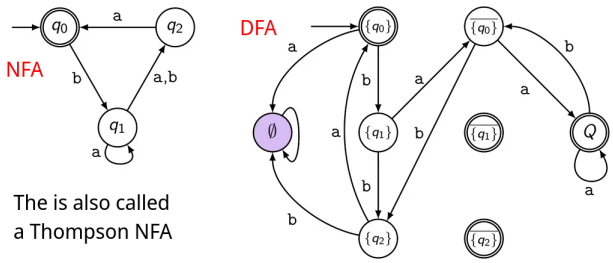
P(Q) is the Potence Quantity!!

Note the P(Q), it means that we have more complex transitions!

no arrow, multiple arrows for a certain character. See b in example

General tipp for NFA, only write what you need to accept the word!

NFA To DFA



$\overline{q}_1$ is the complement Quantity. It means, any state other than q_1

Q is the full Quantity, in here the NFA can be in any state.

$\emptyset$ is the error state.

Formal Definition of the Transition

DFA is marked with ', the regular expression is for NFA

Given δ of an NFA and Transitions $M \subset Q$

$\delta' : Q \times \Sigma \to P(Q) : (M, a) \mapsto \delta'(M, a) = \bigcup_{q \in M} \delta(q, a)$

$Q' = P(Q)$

$\Sigma' = \Sigma$

$q_0' = \{q_0\}$

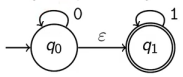
$F' = \{M \in P(Q) | F \cap M \neq \emptyset\}$

M is the union of all possible endstates.

Note: a Thompson NFA needs at least 1 acceptable endstate!

ε Transitions in NFA's

$L = \{0^k 1^l \mid k, l \geq 0\}$



$\delta : Q \times (\Sigma \cup \{\varepsilon\}) \to P(Q)$

This means that ε signifies a transition without using a character!!

Conversion from ε-NFA to regular NFA

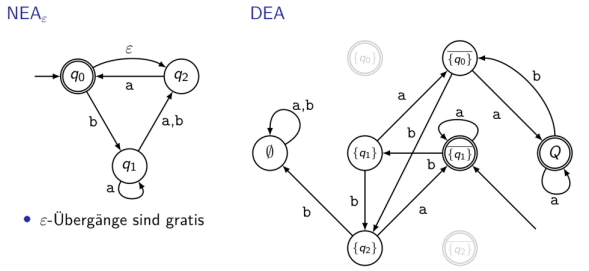
Any ε-NFA can be converted to a regular NFA!

$E(q)$ =Quantity of all ε-Transitions from q

$E(M) = \cup_{q \in M} E(q)$ Quantity of all ε-Transitions

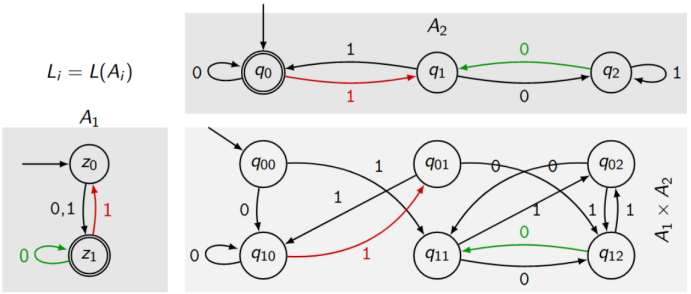
$\delta = Q \times (\Sigma \cap \{\varepsilon\} \to P(Q) : (q, a) \mapsto E(\delta(q, a)))$

ε-NFA to DFA:



Set-Operations with Automatons

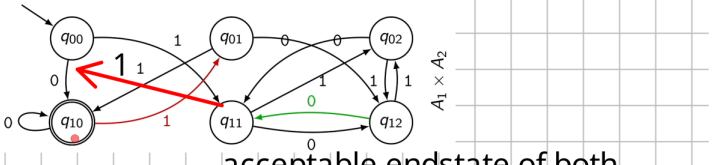
$L_i = L(A_i)$



Intersection $L_1 \cap L_2$

$F = F_1 \times F_2$

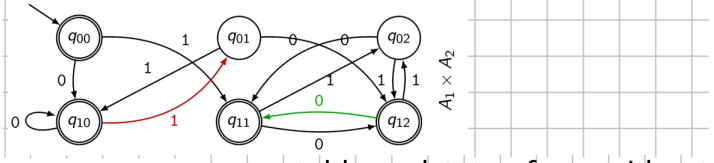
carthesian product -> acceptable endstate which both share



Union $L_1 \cup L_2$

$F = F_1 \times Q_2 \cup Q_1 \times F_2$

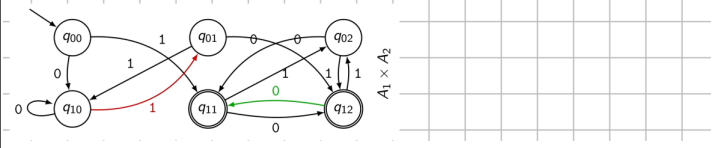
acceptable endstate of both combined



Difference $L_1 \setminus L_2$

$F = F_1 \times (Q_2 \setminus F_2)$

acceptable endstate of one without the ones from the other



Pump-able, but not regular

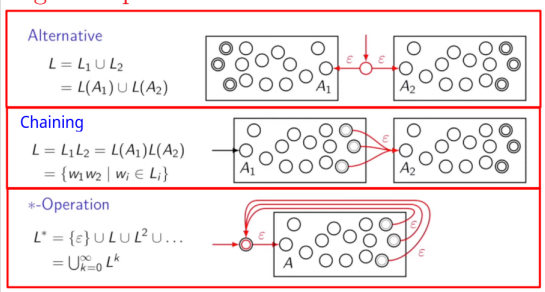
$L = \{a^i b^j c^k \mid i = 0 \vee j = k\} = \underbrace{\{b^j c^k \mid j, k \geq 0\}}_{= L_1} \cup \underbrace{\{a^i b^j c^k \mid i > 0 \wedge j = k\}}_{= L_2}$

The first part L1 is regular, but the second isn't

The only way we can figure that out is by using the Myhill method!

Because L2 is not regular, the composite Language L is not regular

Regular Operations



if or branching statement

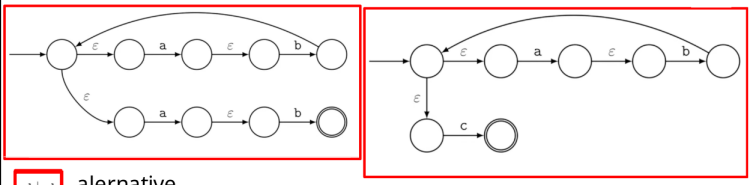
one after another

loop

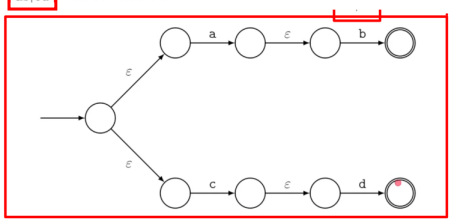
Regular Expressions

$(ab)^+$ do while

$(ab)^*c$ chain



ab|cd alernative



Regular expression

String r to describe a language

L = L(r)

Regular Operations

$L(r_1) \cup L(r_2) = L(r_1 | r_2)$ Alternative

$L(r_1)L(r_2) = L(r_1 r_2)$ Chaining

$L(r_1)^* = L(r_1 *)$ *-Operation

Primitive regular expressions

expressions for Words with length <= 1

$L = L(r)$	r	NEA
$\emptyset$	$\emptyset$	
$\{\varepsilon\}$	ε	
$\{a\}$	a	
$\{o, s, t\}$	[ost]	
$\{a, b, \ldots, s\}$	[a-s]	
Σ	.	

Contractions

$r+ = rr*$ $\quad r? = \varepsilon | r = |r$

$r\{2\} = rr$ $\quad r\{2,3\} = rr|rrr$

$r\{,3\} = |r|rr|rrr$ $\quad r\{3,\} = rrrr*$

Generalized NFA/NEA

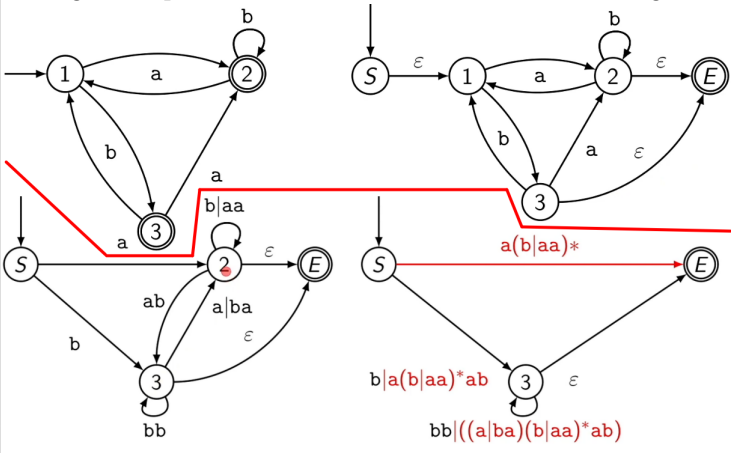
NEA$_\varepsilon$, that has it's transitions described

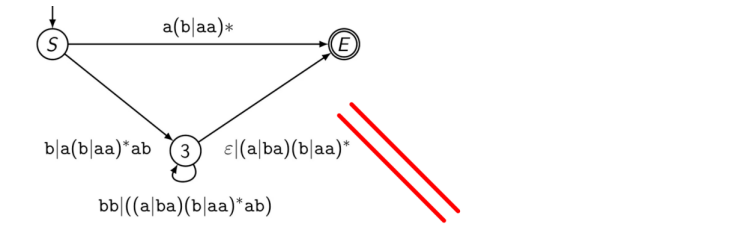
NFA$_\varepsilon$, as regular expressions

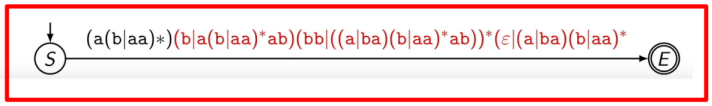
There is a DFA for every regular expression!

VNEA / VNFA to regular expression

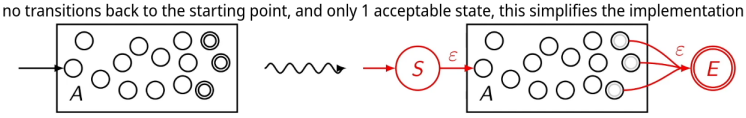
1. add a new start and stop state. This is required as the regular expressions can't have arrow back to the origin.



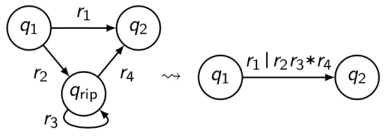




no transitions back to the starting point, and only 1 acceptable state, this simplifies the implementation

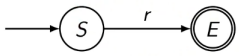


Reduction



Regular Expression

Remove all transitions until you have 1 single regular expression r in A.

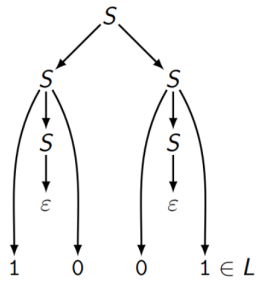


Contextfree-Grammar (CFG) » $G = (V, \Sigma, R, S)$

V: Variables / Non-Terminalsymbols
Σ: Terminalsymbols (Alphabet)
R: Rules of Form $A -> x_1X2...x_n$
with $A \in V, x_i \in V \cup \Sigma$
S: Starvariable
Rule A -> w generated out of uAv
$uwv : uAv \implies uwv$
derivate v from u:
$u \Rightarrow u_1 \Rightarrow u_2 \Rightarrow ... \Rightarrow u_n \Rightarrow v$ or $u \Rightarrow^* v$
Context free grammar generated from G:
$L(G) = \{w \in \Sigma^* | S \Rightarrow^* w\}$

Parse-Tree
$L = \{w \in \{0,1\}^* \; | |w|_0 = |w|_1\}$

with grammar:
$S \rightarrow 0S1 \mid 1S0 \mid SS \mid \varepsilon$



Example:
Grammar $L = \{0^n1^n | n \geq 0\}$
Variables: $V = \{S\}$
Terminalsymbols: $\Sigma = \{0, 1\}$
Rules: $R = \{S \rightarrow \varepsilon | 0S1\}$
Startvariable: S

ε
$01 = 0\varepsilon1$
$0011 = 0\,01\,1$

Contextfree-Grammar Term explanations

Terminalsymbols: Symbols that we can't translate further, aka they are terminal.
CFL = Contextfree-Language
$\Rightarrow^*$ means can be derived from
$uAv \Rightarrow^* uwv$: all words that can be made with uAv
The goal of this grammar is to only have terminal symbols in the end
Aka we remove the variables one by one, via these rules

Definition of Context-free

The context is only based on the input variable.
Take A as a variable, if it can be parsed wihout regard for
what is to the left or right of it, then it has a context free rule
$A \rightarrow$ something
A grammer with only variables like A is considered context free.

Rules without context:
$S \rightarrow A \mid C$
$A \rightarrow a$
$A \rightarrow aAb$
$C \rightarrow bA$

Rules with context:
$S \rightarrow C \mid aC \mid bC$
$aC \rightarrow A$
$bC \rightarrow B$

This means that something like: $S = aAb$
Turns into this: $S \rightarrow aAb \rightarrow aab$
The A can be replaced without regard for context,
S is therefore context free.

Multiple Contextfree-Grammars

Context-free grammars can be interpreted in multiple ways:
This means there is no definite way to interpret something like:
0^n1^n

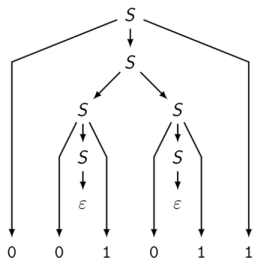
$L = \{w \in \Sigma^* \; | |w|_0 = |w|_1\}$

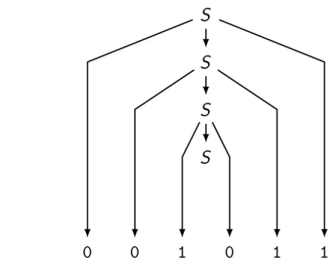
G_1:
$S \rightarrow \varepsilon$
$\rightarrow SS$
$\rightarrow 0S1$
$\rightarrow 1S0$

G_2:
$S \rightarrow SB \mid \varepsilon$
$B \rightarrow N \mid E$
$N \rightarrow NN \mid 0N1 \mid 01 \quad L(N) = \{w \mid w \text{ 0-1- expression}\}$
$E \rightarrow EE \mid 1E0 \mid 10 \quad L(E) = \{w \mid w \text{ 1-0- expression}\}$

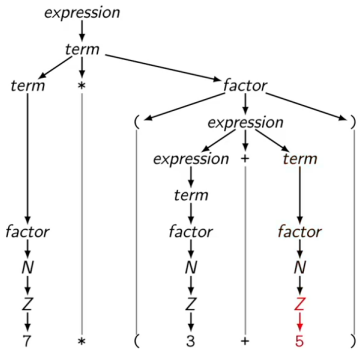
The grammar G_2 gives us a clear Parse-tree, the first one doesn't
The reason for this is the obvious haskell like structure on the right.
The one on the left has no idea of leafs and nodes.
Another example with parse trees:

$S \rightarrow 0S1 \mid 1S0 \mid SS \mid \varepsilon$





$expression \rightarrow expression + term$
$\rightarrow term$
$term \rightarrow term * factor$
$\rightarrow factor$
$factor \rightarrow (\ expression\ )$
$\rightarrow N$
$N \rightarrow NZ$
$\rightarrow Z$
$Z \rightarrow 0|1|2|3|4|5|6|7|8|9$



Facts about Contextfree-Grammar:

- Contextfree-Grammars can't be compared!
- Only the $L_1 \cup L_2$ Quantity operation can be done
  On ANY Contextfree-Grammar
- The amount of steps to derive a word is not always clear
  Some Grammars allow this, some don't
- the class of regular languages is complete with regular operations

Regular Operations on Contextfree-Grammar

Grammar for Regular Operations
L_1 and L_2 context free languages with
Grammars $G_i = (V_i, \Sigma, R_i, S_i)$.
- new starting variable S_0
- $V = V_1 \cup V_2 \cup \{S_0\}$
- extended rules R

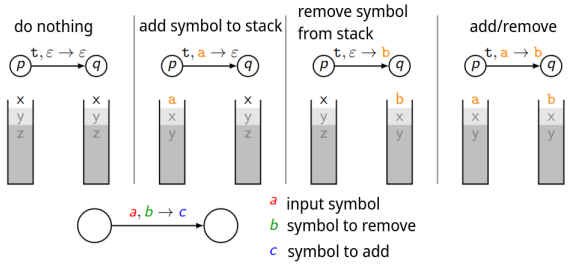
$\Rightarrow G = (V, \Sigma, R, S_0)$

Alternative
Rules for $L_1 \cup L_2$:
$R = R_1 \cup R_2 \cup \{S_0 \rightarrow S_1, S_0 \rightarrow S_2\}$

Chaining
Rules for L_1L_2:
$R = R_1 \cup R_2 \cup \{S_0 \rightarrow S_1S_2\}$

*-Operation
Rules for L_1^*:
$R = R_1 \cup \{S_0 \rightarrow S_0S_1, S_0 \rightarrow \varepsilon\}$

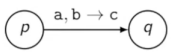
Stack Automaton



depending on the definition, the input can be anything!

Definition

Stack-Automaton $P = (Q, \Sigma, \Gamma, \delta, q_0, F)$ input a, remove b, add c

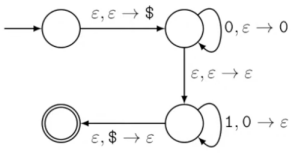


1. Q: States
2. Σ: Input Alphabet
3. Γ: Stack-Alphabet
4. δ: $Q \times \Sigma_\varepsilon \times \Gamma_\varepsilon \to P(Q \times \Gamma_\varepsilon)$
5. $q_0 \in Q$: Start-State
6. $F \subset Q$: Accepting-State

$\Sigma_\varepsilon = \Sigma \cup \{\varepsilon\}$, $\Gamma_\varepsilon = \Gamma \cup \{\varepsilon\}$

Note:
- always an NFA, never deterministic!
- $\Gamma \neq \Sigma$ possible! you can have different symbols in your stack!

Stack Automaton

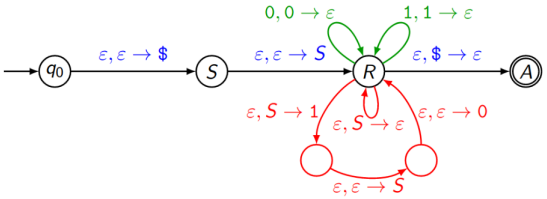


Grammar

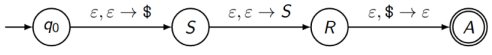
$S \to 0S1$
$\to \varepsilon$

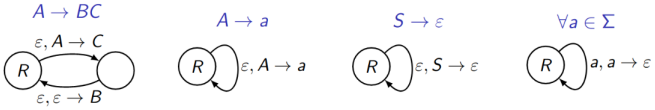
Schema of the Automaton

1. $ and Startvariable S
2. Transition for every Rule
3. Transition for everv Terminalsvmbol



Grammar in CNF, Chomskynormalform -> Stack-Automaton





if L is a Contextfree-language, then an automaton P with L = L(P) exists.
aka this automaton accepts the language L.

Chomsky-Normalform

This is used to make sure, that we don't have needless Variables

1. Unit-rules

$$\left.\begin{array}{l} A \to B \\ B \to A \\ A \to \text{a} \end{array}\right\} \Rightarrow \quad A \to B \to \underbrace{A \to B \to \cdots \to A}_{\text{can be infinite!!}} \to \text{a}$$

2. endless empty Variables

$$\left.\begin{array}{l} A \to ABCDE \mid \text{a} \\ B \to \varepsilon \\ C \to \varepsilon \\ D \to \varepsilon \\ E \to \varepsilon \end{array}\right\} \Rightarrow \quad A \to ABCDE \to ABCD \to ABC \to AB \to A \to \text{a}$$

same here !!

In order to solve this, we need to follow these rules:

1. No Unit Rules: $A \to B$
2. No Rules like: $A \to \varepsilon$ unless necessary: $S \to \varepsilon$
3. Right side has exactly 2 Variables, or 1 terminalsymbol. Everything else if forbidden on the right side.

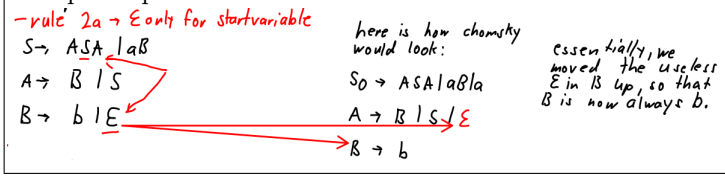
Note this includes ANY other formation, including aB, aa, ABC, aAB

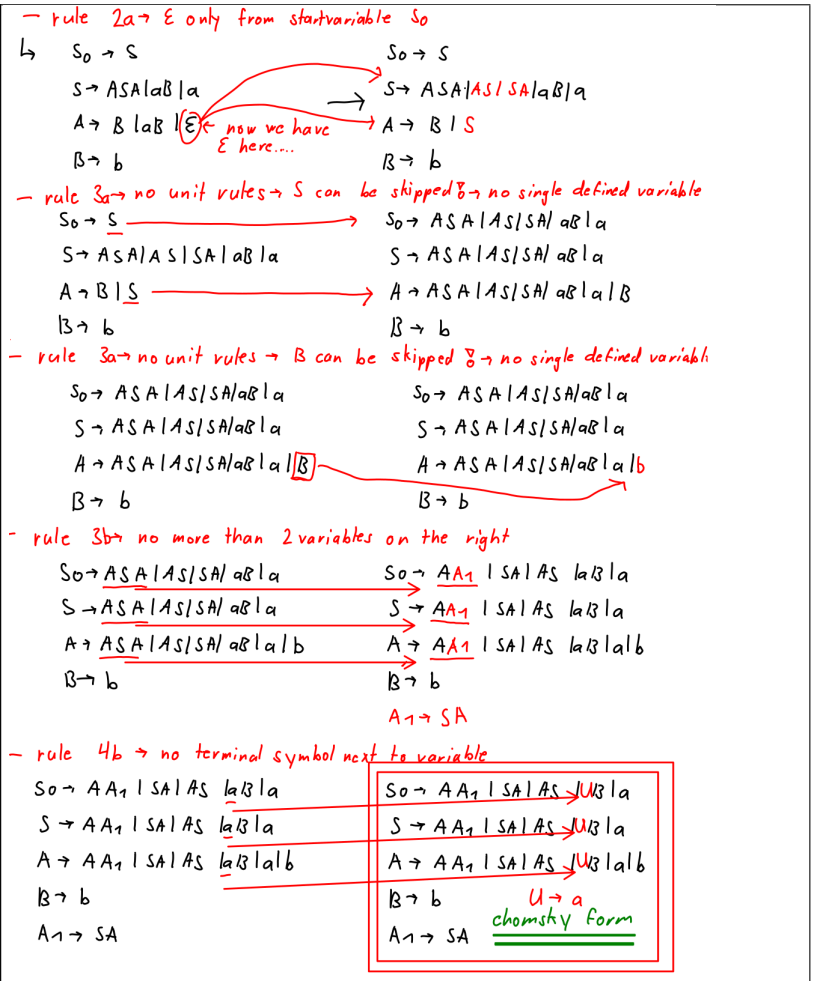
Transforming into Chomsky-Normalform

Transformation into Chomsky-Normalform

1. new starting variable $S_0 \to S$
2. ε- Rules : $\left.\begin{array}{l} A \to \varepsilon \\ B \to AC \end{array}\right\} \Rightarrow$ A is pointless, as it leads to ε. It can simply be emitted. $\Rightarrow \begin{cases} B \to AC \\ \to C \end{cases}$
3. Unit-Rules: $\left.\begin{array}{l} A \to B \\ B \to CD \end{array}\right\} \Rightarrow$ A leads to B, which simply leads to CD. Replace B with CD in A $\Rightarrow \begin{cases} A \to CD \\ B \to CD \end{cases}$
4. Chaining : $A \to u_1 u_2 \ldots u_n$
   $A \to u_1 A_1, A_1 \to u_2 A_2, \ldots, A_{n-2} \to u_{n-1} u_n$ and if u_i == terminalsymbol :
   $A_{i-1} \to U_i A_i, U_i \to u_i$.

In-depth explanation:





Usage of Chomsky-Normalform Grammar G

The derivation of a word $w \in L(G)$ is always possible within $2 * |w| - 1$ Rule applications.
Reason:

1. Apply $|w| - 1$ Rules of form $A \to BC$ to generate a word with length $|w|$ out of S.
2. Apply $|w|$ Rules of form $A \to a$ to generate the word w

This is a total of $2 * |w| - 1$ Rule applications

Deterministic Parsing CYK-Algorithm

Deterministic Parsing is used to check if a word can be derived from either S, aka the entire Grammar $S \Rightarrow^* w$,
or just a variable of it for example A. $S \Rightarrow^* w | A \in V$
This is especially easy with the Chomsky-Normalform!!

Given: Word $w \in \Sigma^*$ Variable $A \in V$ Grammar $G = (V, \Sigma, R, S)$

$A \overset{*}{\Rightarrow} w$ possible??

- $w = \varepsilon$: $A \overset{*}{\Rightarrow} \varepsilon \quad \Leftrightarrow \quad A \to \varepsilon \in R$
- $|w| = 1$: $A \overset{*}{\Rightarrow} w \quad \Leftrightarrow \quad A \to w \in R$
- $|w| > 1$: divide and conquer $A \overset{*}{\Rightarrow} w \quad \Rightarrow \quad \exists \begin{cases} A \to BC & \in R \\ w = w_1 w_2 & w_i \in \Sigma^* \end{cases}$ with $\begin{cases} B \overset{*}{\Rightarrow} w_1 \\ C \overset{*}{\Rightarrow} w_2 \end{cases}$

Wort: ()[()]

$S \to AB \mid CD \mid AT \mid CU \mid SS$
$T \to SB$
$U \to SD$
$A \to ($
$B \to)$
$C \to [$
$D \to]$

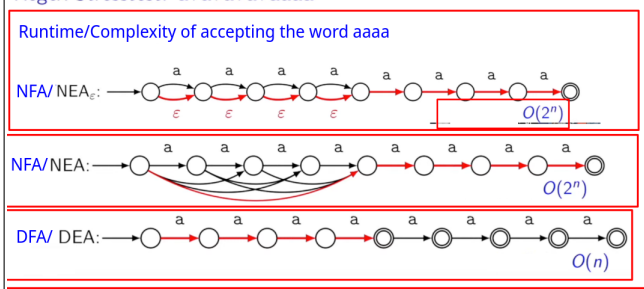
	(	)	[	(	)	]
0	{A}	{B}	{C}	{A}	{B}	{D}
1	{S}	{}	{}	{S}	{}	
2	{}	{}	{}	{U}		
3	{}	{}	{S}			
4	{}	{}				
5	{S}					

okay this works, but what is the catch?
Well $O(n^3)$ ... hahahaa .....

Performance Comparison NFA/DFA

Regex-Stresstest: a?a?a?a?aaaa

Runtime/Complexity of accepting the word aaaa



So what is this? Essentially, because we removed the variables and possibilities, we now have a runtime or complexity of o(n) instead of something with an exponential increase.

Standardization of Stack Automatons (PDA to CFG)

PDA, Push-Down-Automaton = Stack Automaton

The idea is to check whether or not Contextfree-Grammar can be expressed using only regular expressions.

One way of trying this is to create a Stack Automaton, that is based on such expressions.

first need a Variable that handles words from input to stack A_{pq}

Then we need rules for this $A_{pq} \to A_{pr}A_{rq}$

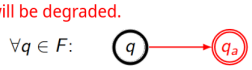
from p to r and from r to q.

Note that A_{pq} signifies the amount of words that lead from p to q

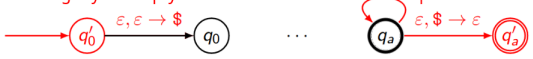
However, this is only the case because it is implied in the name, it is otherwise just a Non-Terminalsymbol!!!

Turning the regular Stack Automaton to a CFG:

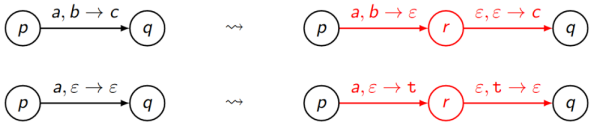
1. only 1 accepting state, q will be degraded. all previous accepting states now lead to qa



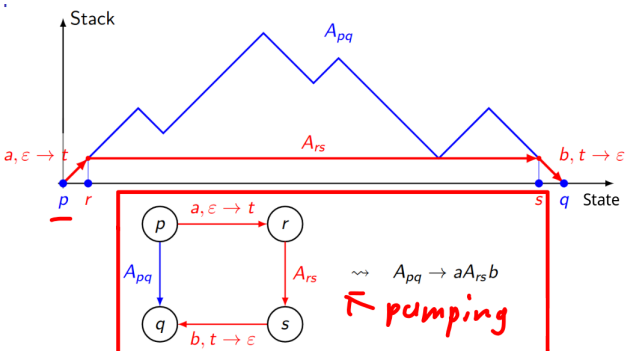
2. on input and output we need to add and remove the $ symbol to signify an empty stack

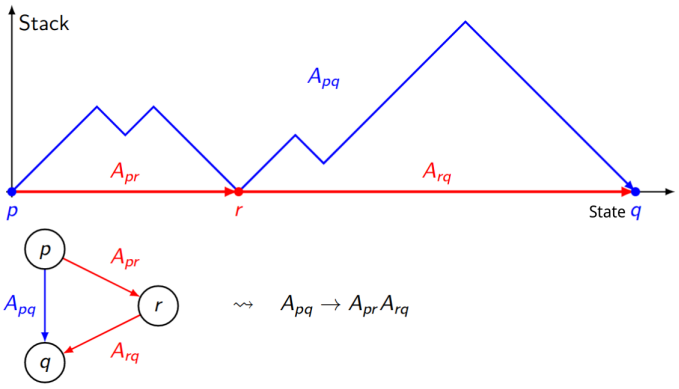


3. now all transitions need to involve the stack



This means that adding or removing need to be different operations.
It also means, that no symbol leads to the stack adding and removing a placeholder.
Here this is t.

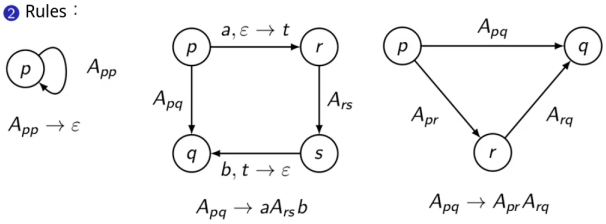




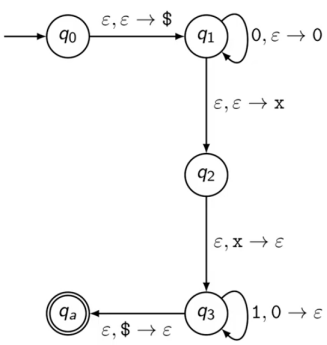
Derived Rules from the Automaton:

standardized Grammar with starting state/accepting state: $q_0 / F = \{q_a\}$.

1. Startvariable: A_{q_0,q_a}
2. Rules :



PDA → CFG: example 1^n0^n



Grammar

$A_{q_0q_a} \to \varepsilon A_{q_1q_3}\varepsilon$
$A_{q_1q_3} \to 0A_{q_1q_3}1$
$\to \varepsilon A_{q_2q_2}\varepsilon$
$A_{q_2q_2} \to \varepsilon$

Simplification, and Chomsky compliant!!

$S \to 0S1$
$\to \varepsilon$

Grammar for: $L = \{0^n1^n \mid n \geq 0\}$

every CFG can be converted to a Stack Automaton.

Regularity Check for CFG / Pumping Lemma for CFG

Grammar G in CNF

$w \in L(G) \quad \Rightarrow \quad S \overset{*}{\Rightarrow} w$

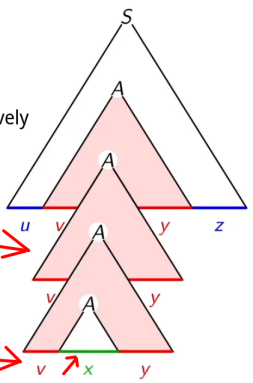
Reusable Variable, RECURSIVE Variable

$|w| \geq N$ considering N big enough, we can recursively use variables.

In this case, the lowest variable is A, so we use this one to recursively call.

$A \overset{*}{\Rightarrow} vxy$

$A \overset{*}{\Rightarrow} x$



Pumping

$A \overset{*}{\Rightarrow} vxy$ instead of simply $A \overset{*}{\Rightarrow} x$

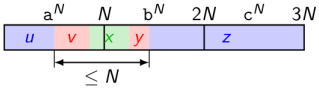
Pumping Lemma for CFL

if L is a CFL, then there is a Number N (Pumping Length) for which:
each word $w \in L$ with $|w| \geq N$
can be divided into $w = uvxyz$

1. $|vy| > 0$
2. $|vxy| \leq N$
3. $uv^kxy^kz \in L \; \forall k \in \mathbb{N}$
5. While pumping, the amount of a and b increases, the amount of c doesn't
   $\Rightarrow uv^kxy^kz \notin L \forall k \neq 1$

Example: $\{a^nb^nc^n \mid n \geq 0\}$

1. Claim : L = CFL
2. Pumping length N
3. Word : $w = a^Nb^Nc^N$
4. Division :



6. Contradiction! L is not a CFL!!

increasing a and b leads to c being smaller!

increasing b and c leads to a being smaller!

even if you increase ALL of them, you still have aa...(your pump)..bb which is NOT correct!

Bakus-Naur Form

-Naur is a machine specification for Rules of a CFG.

- Variables: <variable-name>
- single Symbols: A
- Strings: 'Example'
- Rules: <variable-name> ::= expression

Haskell says hello!

Expressions are a series of variables, single symbols or strings, separated by '|'

```
<expression> ::= <expression> + <term> | <term>
<term>       ::= <term> * <factor> | <factor>
<factor>     ::= ( <expression> ) | <number>
<number>     ::= <number> <digit> | <digit>
<digit>      ::= 0 | 1 | 2 | 3 | 4
                 | 5 | 6 | 7 | 8 | 9
```

Extended-Backus-Naur-Form EBNF

- Literals: "" or ''
- Rules are annotated with = instead of ::=
- Variables without <> and can include symbols
- comma for chaining: term, "*", expression
- semicolon for end, just like cpp: ;
- Comments: (* ... *)
- Optional recursion: {...}
- Option: [] – Groups: (...)
- Exception: - (without, not error handling)

```
expression = expression, "+", term | term ;
term       = term, "*", factor, | factor ;
factor     = "(", expression, ")" | number ;
number     = digit, { digit } ;
digit      = "0" | "1" | "2" | "3" | "4" |
             "5" | "6" | "7" | "8" | "9" ;
```

Both of these Naur implementations do of course resemble this:

$expr \rightarrow term\ expr'$ | $term \rightarrow factor\ term'$ | $factor \rightarrow (\ expr\ )$
$expr' \rightarrow +\ term\ expr'$ | $term' \rightarrow *\ factor\ term'$ | $\rightarrow integer$
$\rightarrow \varepsilon$ | $\rightarrow \varepsilon$ |

One small problem with all this, these specifications change often
So even EBNF is likely soon to be out-of-date....

Note, EBNF does NOT save you from misimplementing things!

```
<expression> ::= <numeric_expression>
...
.
<numeric_expression> ::=
    ( ("-" | "++" | "--") <expression>)
    |( <expression> ("++" | "--" ) )
    |( <expression> ("+" | "+=" | "-" | "-=" | "*" | "*="
        | "/" | "/=" | "%" | "%=") expression).
```

wrong implementation bc lol

Without operator prioritization, + would be calculated before the *.
This means that 5 + 6 * 2 would evaluate to 22 instead of 17.
aka this implementation considers this to be: (5 + 6) * 2
instead of 5 + (6 * 2)

In this picture the Operator-prioritization was not implemented!
This means this implementation is not correct, it calculates trash!

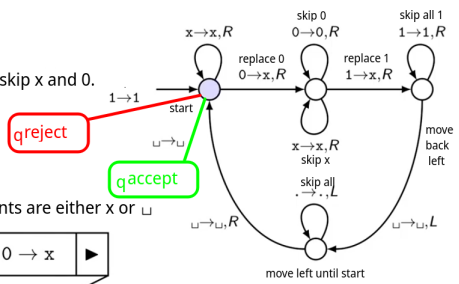
The stack problem

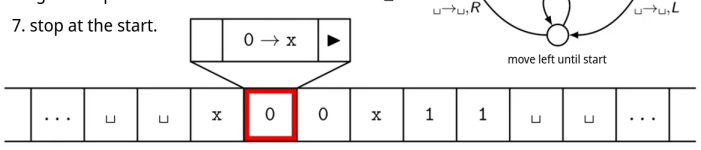
The stack has enabled us to understand the Language 0^n1^n
However the language $a^nb^nc^n$ remains uninterpretable.
This can be solved by extending the stack.
One idea is to double the stack, giving us access to every element
The only problem is that this is a slow and tedious process.
This is where a turing machine comes into play with it's band!!

Turing Machine (TM)

Pseudocode

1. replace 0 with x
2. when moving to the right, skip x and 0.
3. replace the first 1 with x.
4. move to the right until ␣
5. move to the left until ␣
6. go to step 1 until all elements are either x or ␣
7. stop at the start.

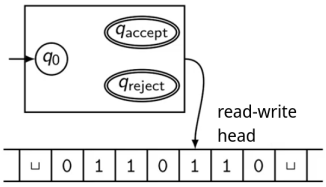




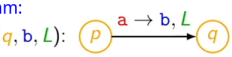
Definition Of Deterministic Turing Machine

Turing Machine
$M = (Q, \Sigma, \Gamma, \delta, q_0, q_{\text{accept}}, q_{reject})$

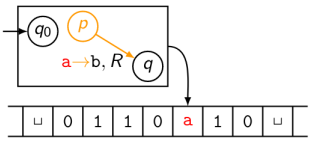
- Q : States
- Σ : Alphabet
- Γ : Band-Alphabet: $\sqcup \in \Gamma \backslash \Sigma$
- δ : $Q \times \Gamma \rightarrow Q \times \Gamma \times \{L, R\}$
- $q_0 \in Q$: Startstate
- $q_{accept} \in Q$: accepting state
- $q_{reject} \in Q$: rejecting state



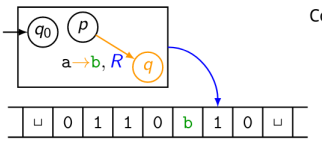
- Transition possible if a under read-write head
- Current field of read-write head is overwritten with b
- movement head: left L, right R

State Diagram:
$\delta(p, a) = (q, b, L)$: 

Before



After



Programstructure

- Input-Word on Band
- Read-Write Head positioned on 1. element
- start Machine, complete t(w) transitions
- Machine stops in qaccept or qreject

Complexity: $t(w)$, $t(n) = \max\{t(w) \mid |w| \leq n\}$

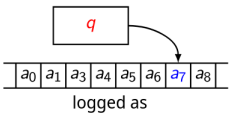
of a TM recognized language

M a Turing Machine

$L(M) = \{w \in \Sigma^* \mid M \text{ accepts } w\}$

Calculation History

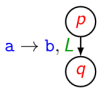
Notation for Machine State



logged as

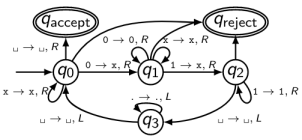
$a_0 a_1 a_2 a_3 a_4 a_5 a_6 q a_7 a_8$

Transition



$a_1\ a_2\ a_3\ p\ a\ a_4\ a_5$
$a_1\ a_2\ q\ a_3\ b\ a_4\ a_5$

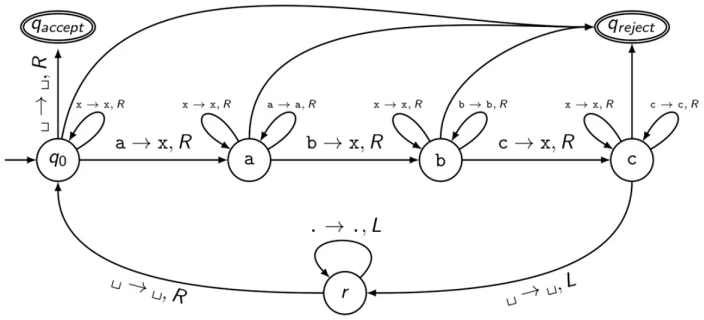
Protocoll of a Calculation



␣ q0 0 0 1 1 ␣
␣ x q1 0 1 1 ␣
␣ x 0 q1 1 1 ␣
␣ x 0 x q2 1 ␣
␣ x 0 x 1 q2 ␣
␣ x 0 x q3 1 ␣
␣ x 0 q3 x 1 ␣
␣ x q3 0 x 1 ␣
␣ q3 x 0 x 1 ␣
...

$a^nb^nc^n$

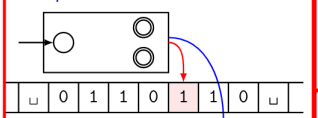
Before we asked the quetion of how we could create a Turing Machine, that understands the language $a^nb^nc^n$
Is this already possible ? YES!



Turing Machine Variants

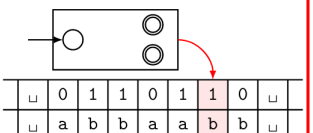
different Alphabet
Γ Binary encoded Bandalphabet
$\Gamma_0 = \{0, 1\}$
Complexity $O(t(n))$

multiple Bands



Complexity $O(t(n)^2)$

Multirow Turing Machine



Complexity $O(t(n))$

non-deterministic TM
δ: $Q \times \Gamma \rightarrow P(Q \times \Gamma \times \{L, R\})$
Complexity $2^{O(t(n))}$

Nondeterministic TM

Transitioning Function

on each step max N different possibilities

$\delta: Q \times \Gamma \to P(Q \times \Gamma \times \{L, R\})$

$\Rightarrow$ multiple directions of movement

accept Word

$w \in L(M)$

a word is acceptable if there is at least one possible way to accept it. The rest do not matter!

Simulation Idea:

Try all posibilities $\max N^{t(n)}$

Simulation on standard TM

use 3 bands

1. Working Band
2. copy of w (word)
3. A list of all sequences of possibilities

Simulation

1. copy word to band 1
2. execute TM on band 1 with possibilities of band 3
3. start again at step 1 and choose the next possibility from band 3

Complexity: $O(N^{t(n)}) = 2^{O(t(n))}$

Counter for TM

A language of a Turing Machine is countable

It essentially allows you to see all possible words.

Definition

A counter is a TM which allows you to print words

recursively countable Language

L is recursive countable if there is a counter that counts it.

Countable Language == Turing Language

Definitions for Computability (Entscheidbarkeit)

Decider

Definition

A Decider is a Turing-Machine that stops on ANY input

Definition

A Language L is computable, if there is a Decider M with $L = L(M)$.

M decides L.

Language-Problem

Every Problem P can be turned into a Language Problem

$$L_P = \left\{ w \in \Sigma^* \;\middle|\; \begin{array}{l} w \text{ is the solution of} \\ \text{Problem P} \end{array} \right\}$$

Examples:

Empty-Problem :

$$E_{\text{DEA}} = \left\{ \langle A \rangle \;\middle|\; \begin{array}{l} A \text{ a DFA, and} \\ L(A) = \emptyset \end{array} \right\}$$

Equality Problem :

$$EQ_{\text{CFG}} = \left\{ \langle G_1, G_2 \rangle \;\middle|\; \begin{array}{l} G_i \text{ CFGs and} \\ L(G_1) = L(G_2) \end{array} \right\}$$

Acceptance Problem :

$$A_{\text{DEA}} = \left\{ \langle A, w \rangle \;\middle|\; \begin{array}{l} A \text{ a DFA that} \\ \text{accepts } w \end{array} \right\}$$

Halting Problem :

$$HALT_{\text{TM}} = \left\{ \langle M, w \rangle \;\middle|\; \begin{array}{l} M \text{ stops at} \\ \text{Input } w \end{array} \right\}$$

There are clear definite problems that we can solve:

Language Problem Calc.

find the solution of the quadratic formula

$x^2 - x - 1 = 0$

a b c and x are decimals, aka numbers

$w = \texttt{a=}a\texttt{,b=}b\texttt{,c=}c\texttt{,x=}x$

Computability problem: is this a solution? Yes? No?

$\texttt{a=1,b=-1,c=-1,x=3} \in L$?

ε acceptance Problem

can the Machine accept the empty word?

$L = \{\langle A \rangle \mid \varepsilon \in L(A)\}$

Compute-Algorithm

1. Turn A into a DFA
2. is the startstate a accepting state?

Language Problem 1 word

can the Machine accept the word w?

$L = \{\langle A, w \rangle \mid w \in L(A)\}$

Compute-Algorithm

1. turn A into a DFA
2. Simulate A with a Turing Machine
3. does the Turing Machine stop in qaccept?

Given n calculate the first 10 decimals of its square root

L is the language of strings of form n,x where n is a decimalform of a natural number and x is the first 10 decimals of the root

Computability is this a solution? Yes? No?

$\texttt{2,1.414213562} \in L$?

or $\langle \text{n,x} \rangle \longrightarrow \langle A \rangle, \in L$?

Do Machine A1 and A2 accept the same language?

$L(A_1) = L(A_2)$?

$L = \{\langle A_1, A_2 \rangle \mid L(A_1) = L(A_2)\}$

Computability

1. Turn A1 into a minimal Automaton A'1.
2. Turn A2 into a minimal Automaton A'2.
3. is A'1 = A'2?

Can the word w be produced from the grammar G?

$L = \{\langle G, w \rangle \mid w \in L(G)\}$

Computability

1. Check if <G> is actually a Grammar
2. Bring Grammar into CNF (Chomsky)
3. check with the CYK algorithm if w can be produced out of G.

And there are some that we can't solve:

Theorem (Alan Turing)

A_{TM} is not computable

Proof:

Generate from a decider H for A_{TM} a machine D with Input $\langle M \rangle$

1. let H on Input $\langle M, \langle M \rangle \rangle$
2. if H accepts : q_{reject}
3. if H rejects : q_{accept}

Now use D on $\langle D \rangle$

$D(\langle D \rangle)$ accepts $\Leftrightarrow$ D rejects $\langle D \rangle$

$D(\langle D \rangle)$ rejects $\Leftrightarrow$ D accepts $\langle D \rangle$

Contradiction!

Special-Halting Problem

$$HALT\varepsilon_{\text{TM}} = \left\{ \langle M \rangle \;\middle|\; \begin{array}{l} M \text{ is a turing machine} \\ \text{and } M \text{ stops on empty} \\ \text{band} \end{array} \right\}$$

is not computable

Halting Problem

$$HALT_{\text{TM}} = \left\{ \langle M, w \rangle \;\middle|\; \begin{array}{l} M \text{ is a turing machine} \\ \text{and } M \text{ stops on input w} \end{array} \right\}$$

is not computable

Halting Problem

The idea is that we have a Decider H and a program D that always returns the op

This D encompasses H, which means we give input to D, which gives the input to

Then H returns and D inverts it, as described above.

The contradictin arises when you give D its own code as the input.

D passes this on and returns an answer, but D inverts it, aka does the opposite

This means H was wrong, aka the decider has lied, D didn't do what it said!

Reduction

Reductionprojection

solveable projection $f: \Sigma^* \to \Sigma^*$

$w \in A \quad \Leftrightarrow \quad f(w) \in B$

Notation: $f: A \leq B$, "A easier than B"

Computability

B computable , $f: A \leq B \Rightarrow A$ computable

Proof

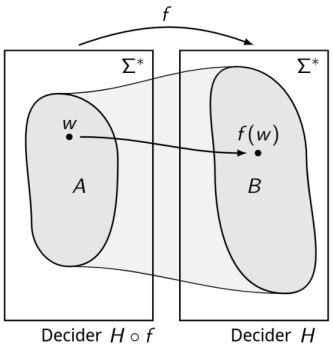
H is a decider for B, then is $H \circ f$ a decider for A.

Conclusion

A not computable , $A \leq B \Rightarrow B$ not computable

or for humans: if B is computable then so is A, as A is the easier problem than B.

If B is NOT computable then A is NOT computable.



Reduction for: $A_{\text{TM}} \leq HALT\varepsilon_{\text{TM}}$

A_{TM} is not computable

Does the machine M accept the word w?

There is no Decider

$\langle M, w \rangle \quad \mapsto$ Program S

1. run M on inputword w
2. M stops on qaccept: accept
3. M stops on qreject: endless loop

M accepts $w \quad \Leftrightarrow \quad S$ stops

is $HALT\varepsilon_{\text{TM}}$ computable?

Does the machine stop?

Decider for A_{TM}

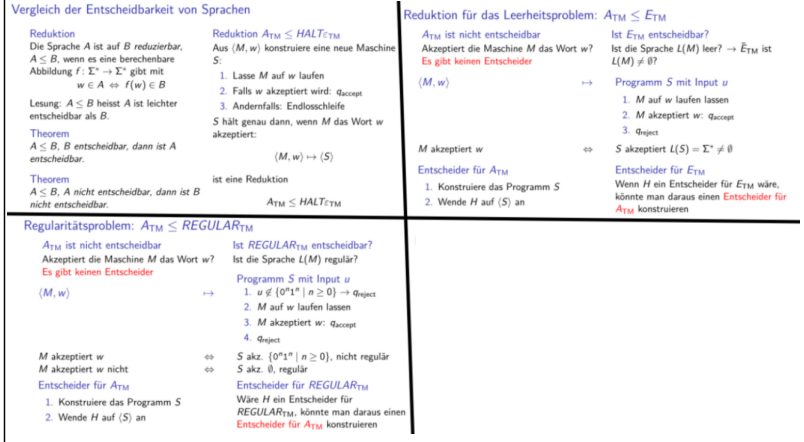
1. Create program S
2. Use H on $\langle S \rangle$

Decider for $HALT\varepsilon_{\text{TM}}$

If H is a decider for $HALT\varepsilon_{\text{TM}}$ then we could create a decider for A_{TM}

The same strategy can be used on many more examples.

Due to time constraints I did not translate these:



The law of Rice

Law of Rice

Properties of a language
(properties of a turing language)

REGULAR	$L(M)$ is regular
E	$L(M)$ is empty

Satz von Rice
if P is a nontrivial property of a turing language, then

$P_{\text{TM}} = \{\langle M \rangle \mid L(M) \text{ has property } P\}$
not computable

Definition
A property is nontrivial is there are 2 Turing machines that handle the same language, with one of them having the property and the other not.

$L(M_1)$ has P
$L(M_2)$ doesn't have P

Conclusion:
It is not possible to see if a program has a nontrivial property or not!

Reduction P_{TM}: $A_{\text{TM}} \leq P_{\text{TM}}$

A_{TM} is not computable
Does the machine M accept the word w?

is P_{TM} computable?
Does the language L(M) have the property P?

assumptions
- Σ^* doesn't have property P
- Testprogram T: $L(T)$ doesn't have P

$\langle M, w \rangle \quad \mapsto$ Programm S mit Input u

1. Program T accepts u: q_{accept}
2. Run M on w
3. M accepts w: qaccept
4. q_{reject}

M accepts $w \quad \Leftrightarrow \quad S$ accepts Σ^*, has P
M rejects $w \quad \Leftrightarrow \quad S$ accepts $L(T)$, doesn't have P

Overview of Computability

Problem	Word	Prerequisite		Computing algorithm / Reason
E_{DEA}	$\langle A \rangle$	$L(A) = \emptyset$	yes	minimal-automaton doesn't have accept state
E_{CFG}	$\langle G \rangle$	$L(G) = \emptyset$	yes	Chomsky-Normalform
E_{TM}	$\langle M \rangle$	$L(M) = \emptyset$	no	
EQ_{DEA}	$\langle A_1, A_2 \rangle$	$L(A_1) = L(A_2)$	yes	comparison of minimal Automatons
EQ_{CFG}	$\langle G_1, G_2 \rangle$	$L(G_1) = L(G_2)$	no	
EQ_{TM}	$\langle M_1, M_2 \rangle$	$L(M_1) = L(M_2)$	no	
A_{DEA}	$\langle A, w \rangle$	$w \in L(A)$	yes	Regex-Engines simulate any DEAs with any input word w
A_{CFG}	$\langle G, w \rangle$	$w \in L(G)$	yes	Cocke-Younger-Kasami algorithm
A_{TM}	$\langle M, w \rangle$	$w \in L(M)$	no	Halting-Problem