

## Der Satz von Cook

Prof. Dr. David Sabel

LFE Theoretische Informatik



Letzte Änderung der Folien: 19. Juli 2022

## SAT-Problem

- Wir brauchen ein erstes Problem, dessen  $\mathcal{NP}$ -Vollständigkeit wir per Hand nachweisen müssen.
- Dafür nehmen wir das SAT-Problem.

### Definition (SAT-Problem)

Das **Erfüllbarkeitsproblem der Aussagenlogik** (kurz **SAT**) ist:

gegeben: Eine Aussagenlogische Formel  $F$

gefragt: Ist  $F$  erfüllbar, d.h. gibt es eine erfüllende Belegung der Variablen mit den Wahrheitswerten 0 und 1, sodass  $F$  den Wert 1 erhält.

Als formale Sprache:

$SAT = \{code(F) \in \Sigma^* \mid F \text{ ist erfüllbare Formel der Aussagenlogik}\}$

## Wiederholung: NP-Vollständigkeit

### Definition ( $\mathcal{NP}$ -Vollständigkeit)

Eine Sprache  $L$  heißt  **$\mathcal{NP}$ -vollständig**, wenn gilt

- 1  $L \in \mathcal{NP}$  und
- 2  $L$  ist  **$\mathcal{NP}$ -schwer** (manchmal auch  $\mathcal{NP}$ -hart genannt):  
Für alle  $L' \in \mathcal{NP}$  gilt  $L' \leq_p L$

Beweistechnik für  $\mathcal{NP}$ -Schwere:

- Zeige  $L_0 \leq_p L$  für ein bekanntes  $\mathcal{NP}$ -vollständiges Problem.
- Dann folgt die  $\mathcal{NP}$ -Schwere von  $L$

## SAT $\in \mathcal{NP}$

### Lemma

SAT  $\in \mathcal{NP}$

Beweis:

- $code(F)$  Eingabe einer NTM  $M$
- $M$  berechnet, welche Variablen in  $F$  vorkommen. Sei dies  $\{x_1, \dots, x_n\}$ .
- $M$  verwendet Nichtdeterminismus, um Belegung zu „raten“:

$$I : \{x_1, \dots, x_n\} \rightarrow \{0, 1\}$$

- Jede der  $2^n$  nicht-deterministischen Berechnungen berechnet Wert von  $I(F)$
- Akzeptanz bei 1, sonst verwerfen.
- Da jede Belegung überprüft wird, gilt:

$M$  akzeptiert eine Formel  $F$  g.d.w.  $F \in SAT$

- Jeder Berechnungspfad von  $M$  läuft in Polynomialzeit in  $|code(F)|$ , da Anzahl der Variablen durch die Eingabegröße beschränkt ist.

## $\mathcal{NP}$ = Polynomiell verifizierbar

- Nachweis, dass Sprache in  $\mathcal{NP}$  liegt, geht oft so wie bei SAT
- Verwende Nichtdeterminismus, um potentielle Lösung zu raten
- Zeige, dass eine Lösung in Polynomialzeit verifiziert werden kann

## $\mathcal{NP}$ -Schwere von SAT: Ideen

- Wir müssen zeigen:

$$L \leq_p \text{SAT für alle } L \in \mathcal{NP}$$

- Da  $L \in \mathcal{NP}$ , gibt es polynomiell zeitbeschränkte NTM  $M$ , die  $L$  akzeptiert.
- Für Wort  $w$  erstelle Formel  $F$  (in determ. Polynomialzeit), sodass gilt:  
 $F$  erfüllbar g.d.w.  $M$  akzeptiert  $w$
- Da Laufzeit polynomiell beschränkt, beschränkt dies auch die Größe der Formel und wir haben  $L \leq_p \text{SAT}$

## Hilfssatz für den $\mathcal{NP}$ -Schwere Beweis von SAT

### Lemma

Für aussagenlogische Variablen  $\{x_1, \dots, x_n\}$  gibt es eine aussagenlogische Formel  $exactlyOne(x_1, \dots, x_n)$ , sodass

$$I(exactlyOne(x_1, \dots, x_n)) = 1 \quad \text{g.d.w.} \quad I \text{ setzt genau eine der Variablen } x_i \text{ auf 1} \\ \text{und alle anderen auf 0}$$

Dabei ist die Größe der Formel  $exactlyOne(x_1, \dots, x_n)$  in  $O(n^2)$ .

Beweis:

$$exactlyOne(x_1, \dots, x_n) := (x_1 \vee \dots \vee x_n) \wedge \bigwedge_{1 \leq i < j \leq n} \neg(x_i \wedge x_j)$$

- $(x_1 \vee \dots \vee x_n)$  sichert  $atLeastOne(x_1, \dots, x_n)$  zu
- $\bigwedge_{1 \leq i < j \leq n} \neg(x_i \wedge x_j)$  sichert  $atMostOne(x_1, \dots, x_n)$  zu

## SAT ist $\mathcal{NP}$ -schwer (1)

### Lemma

SAT ist  $\mathcal{NP}$ -schwer.

Beweis: Wir müssen zeigen:  $L \leq_p \text{SAT}$

- Sei  $L \in \mathcal{NP}$  beliebig.
- Sei  $M$  die NTM mit  $L(M) = L$  und  $ntime_M(w) \leq p(|w|)$ .
- Sei  $w$  eine Eingabe für  $M$ .

Ziel: Konstruiere aussagenlogische Formel  $F$ , sodass gilt

$$w \in L \iff F \text{ ist erfüllbar}$$

Dabei muss  $F$  in Polynomialzeit konstruierbar sein.

D.h. wir geben eine in Polynomialzeit berechenbare Funktion  $f(w)$  an, sodass  $w \in L \iff f(w) \in \text{SAT}$ .

## SAT ist $\mathcal{NP}$ -schwer (2)

Eingabe:	$w = a_1 \cdots a_n \in \Sigma^*$
Bandalphabet:	$\Gamma = \{b_1, \dots, b_l\}$
Zustände:	$Z = \{z_0, \dots, z_k\}$
Startzustand:	$z_0$

### Aussagenlogische Variablen in der Formel $F$ :

Variable	Index-Bereich	Bedeutung
$State_{t,z}$	$t = 0, 1, \dots, p(n)$ $z = z_0, \dots, z_k$	$State_{t,z} = 1$ g.d.w. nach $t$ Schritten ist $M$ im Zustand $z$ .
$Post_{t,i}$	$t = 0, 1, \dots, p(n)$ $i = -p(n), \dots, p(n)$	$Post_{t,i} = 1$ g.d.w. nach $t$ Schritten ist der Schreib-Lesekopf auf Position $i$
$Tape_{t,i,b}$	$t = 0, 1, \dots, p(n)$ $i = -p(n), \dots, p(n)$ $b = b_1, \dots, b_l$	$Tape_{t,i,b} = 1$ g.d.w. nach $t$ Schritten steht in Position $i$ das Zeichen $b$

- Bandpositionen: Position 0 am Anfang
- Bereiche reichen aus, da die TM nicht mehr als  $p(n)$  Schritte macht

## SAT ist $\mathcal{NP}$ -schwer (3)

Aufbau der Formel  $F$ :

$$F = Rand \wedge Anfang \wedge Transition \wedge Ende$$

- *Rand*: Randbedingungen
- *Anfang*: Anfangsbedingungen
- *Transition*: Bedingungen für den Zustandsübergang
- *Ende*: Endbedingung

## SAT ist $\mathcal{NP}$ -schwer (4)

### Randbedingungen:

Zu jedem Zeitpunkt  $t$ :

- ... ist  $M$  in genau einem Zustand  $z$ :

$$\bigwedge_{t \in \{0, \dots, p(n)\}} \text{exactlyOne}(State_{t,z_0}, \dots, State_{t,z_k})$$

- ... ist der Kopf von  $M$  in genau einer Position auf dem Band:

$$\bigwedge_{t \in \{0, \dots, p(n)\}} \text{exactlyOne}(Post_{t,-p(n)}, \dots, Post_{t,p(n)})$$

- ... befindet sich in jeder Bandzelle genau ein Symbol aus  $\Gamma$ :

$$\bigwedge_{t \in \{0, \dots, p(n)\}} \bigwedge_{i \in \{-p(n), \dots, p(n)\}} \text{exactlyOne}(Tape_{t,i,b_1}, \dots, Tape_{t,i,b_l})$$

Daher:

$$Rand := \bigwedge_{t \in \{0, \dots, p(n)\}} \left( \begin{array}{l} \text{exactlyOne}(State_{t,z_0}, \dots, State_{t,z_k}) \\ \wedge \text{exactlyOne}(Post_{t,-p(n)}, \dots, Post_{t,p(n)}) \\ \wedge \bigwedge_{i \in \{-p(n), \dots, p(n)\}} \text{exactlyOne}(Tape_{t,i,b_1}, \dots, Tape_{t,i,b_l}) \end{array} \right)$$

## SAT ist $\mathcal{NP}$ -schwer (5)

### Anfangsbedingungen:

Fixieren die Bedingungen zum Zeitpunkt  $t = 0$ :

- $M$  ist im Startzustand:  $State_{0,z_0}$
- Der Schreib-Lesekopf ist auf Position 0:  $Pos_{0,0}$
- Die Eingabe  $w = a_1 \cdots a_n$  steht auf dem Band und alle anderen Zellen enthalten das Blank-Symbol.

$$\left( \bigwedge_{i \in \{0, \dots, n-1\}} Tape_{0,i,a_{i+1}} \right) \wedge \left( \bigwedge_{i \in \{-p(n), \dots, -1\}} Tape_{0,i,\square} \right) \wedge \left( \bigwedge_{i \in \{n, \dots, p(n)\}} Tape_{0,i,\square} \right).$$

Daher:

$$Anfang := State_{0,z_0} \wedge Pos_{0,0} \wedge \left( \bigwedge_{i \in \{0, \dots, n-1\}} Tape_{0,i,a_{i+1}} \right) \wedge \left( \bigwedge_{i \in \{-p(n), \dots, -1\}} Tape_{0,i,\square} \right) \wedge \left( \bigwedge_{i \in \{n, \dots, p(n)\}} Tape_{0,i,\square} \right)$$

## SAT ist $\mathcal{NP}$ -schwer (6)

### Transitionsbedingungen:

- Für Übergang von  $t$  zu  $t + 1$ : Zustand, Bandinhalt, Position ändern.  
Sei  $dir(N) = 0, dir(L) = -1, dir(R) = 1$ .

$$\bigwedge_{\substack{t \in \{0, \dots, p(n)-1\}, z \in Z, \\ i \in \{-p(n)+1, \dots, p(n)-1\}, b \in \Gamma}} \left( \left( State_{t,z} \wedge Pos_{t,i} \wedge Tape_{t,i,b} \right) \implies \bigvee_{(z',b',y) \in \delta(z,b)} \left( State_{t+1,z'} \wedge Pos_{t+1,i+dir(y)} \wedge Tape_{t+1,i,b'} \right) \right)$$

- Zellen auf denen der Kopf nicht steht, bleiben unverändert.

$$\bigwedge_{\substack{t \in \{0, \dots, p(n)-1\}, \\ i \in \{-p(n), \dots, p(n)\}, \\ b \in \Gamma}} \left( (\neg Pos_{t,i} \wedge Tape_{t,i,b}) \implies Tape_{t+1,i,b} \right)$$

## SAT ist $\mathcal{NP}$ -schwer (7)

### Transitionsbedingungen:

Ergibt zusammen:

$$\begin{aligned} \text{Transition} := & \left( \bigwedge_{\substack{t \in \{0, \dots, p(n)-1\}, \\ z \in Z, \\ i \in \{-p(n)+1, \dots, p(n)-1\}, \\ b \in \Gamma}} \left( \left( State_{t,z} \wedge Pos_{t,i} \wedge Tape_{t,i,b} \right) \implies \bigvee_{(z',b',y) \in \delta(z,b)} \left( State_{t+1,z'} \wedge Pos_{t+1,i+dir(y)} \wedge Tape_{t+1,i,b'} \right) \right) \right) \\ & \bigwedge_{\substack{t \in \{0, \dots, p(n)-1\}, \\ i \in \{-p(n), \dots, p(n)\}, \\ b \in \Gamma}} \left( (\neg Pos_{t,i} \wedge Tape_{t,i,b}) \implies Tape_{t+1,i,b} \right) \end{aligned}$$

## SAT ist $\mathcal{NP}$ -schwer (8)

**Endbedingung:** Ein akzeptierender Zustand wird erreicht:

$$\text{Ende} := \bigvee_{z \in E, t \in \{0, \dots, p(n)\}} State_{t,z}$$

## SAT ist $\mathcal{NP}$ -schwer (9)

$$\begin{aligned} F = & \bigwedge_{t \in \{0, \dots, p(n)\}} \left( \begin{array}{l} \text{exactlyOne}(State_{t,z_0}, \dots, State_{t,z_k}) \\ \wedge \text{exactlyOne}(Pos_{t,-p(n)}, \dots, Pos_{t,p(n)}) \\ \wedge \bigwedge_{i \in \{-p(n), \dots, p(n)\}} \text{exactlyOne}(Tape_{t,i,b_1}, \dots, Tape_{t,i,b_l}) \end{array} \right) \\ & \wedge State_{0,z_0} \wedge Pos_{0,0} \wedge \left( \bigwedge_{i \in \{0, \dots, n-1\}} Tape_{0,i,a_{i+1}} \right) \wedge \left( \bigwedge_{i \in \{-p(n), \dots, -1\}} Tape_{0,i,\square} \right) \wedge \left( \bigwedge_{i \in \{n, \dots, p(n)\}} Tape_{0,i,\square} \right) \\ & \bigwedge_{\substack{t \in \{0, \dots, p(n)-1\}, \\ z \in Z, \\ i \in \{-p(n)+1, \dots, p(n)-1\}, \\ b \in \Gamma}} \left( \left( State_{t,z} \wedge Pos_{t,i} \wedge Tape_{t,i,b} \right) \implies \bigvee_{(z',b',y) \in \delta(z,b)} \left( State_{t+1,z'} \wedge Pos_{t+1,i+dir(y)} \wedge Tape_{t+1,i,b'} \right) \right) \\ & \bigwedge_{\substack{t \in \{0, \dots, p(n)-1\}, \\ i \in \{-p(n), \dots, p(n)\}, \\ b \in \Gamma}} \left( (\neg Pos_{t,i} \wedge Tape_{t,i,b}) \implies Tape_{t+1,i,b} \right) \\ & \bigwedge_{z \in E, t \in \{0, \dots, p(n)\}} State_{t,z} \end{aligned}$$

## SAT ist $\mathcal{NP}$ -schwer (10)

- Wenn  $w \in L$ , dann  $z_0 w \vdash_M^r uz_e v$  mit  $z_e \in E$  und  $r \leq p(n)$ .
- Lauf liefert Belegung  $I$  der Variablen von  $F$ , sodass  $I(F) = 1$ :
  - In *Rand* belege die besuchten Zustände, Positionen und Bandinhalte mit 1. Falls die Folge nach  $t_e < p(n)$  Schritten endet, so setze die Variablen für  $t > t_e$  auf die Werte für den Zeitpunkt  $t_e$ .
  - Lauf liefert Belegung der Variablen in *Anfang*, so dass sie zur Anfangskonfiguration passen
  - Für *Transition* setze  $I(\text{State}_{t,z}) = 1, I(\text{Pos}_{t,i}) = 1, \text{Tape}_{t,i,b}$  entsprechend der besuchten Zustände und alle anderen auf 0. Das macht  $I(\text{Transition}) = 1$  und Variable  $\text{State}_{t,z_e}$  für  $z_e \in E$  wird dadurch auf 1 gesetzt.
- Daher  $I(F) = 1$

## SAT ist $\mathcal{NP}$ -schwer (11)

- Umgekehrt: Wenn es eine erfüllende Belegung  $I$  gibt mit  $I(F) = 1$ , dann kann daraus ein akzeptierender Lauf für die TM auf Eingabe  $w$  konstruiert werden.
- Damit gilt  $w \in L \iff F$  ist erfüllbar.
- $F$  kann in (deterministischer) Polynomialzeit berechnet werden:

Größe von  $F$ : (Anzahl an Variablenvorkommen):

Subformel	Größe
<i>Rand</i> :	$O(p(n)^3)$
<i>Anfang</i> :	$O(p(n))$
<i>Transition</i> :	$O(p(n)^2)$
<i>Ende</i> :	$O(p(n))$
$F$	$O(p(n)^3)$

## Satz von Cook

Insgesamt haben wir damit gezeigt:

### Satz von Cook

Das Erfüllbarkeitsproblem der Aussagenlogik ist  $\mathcal{NP}$ -vollständig.