SAT Solving

Skript zur Vorlesung im WS 2023/24
gehalten von Dr. Jan Johannsen

Version vom 21.03.2023

1 Einführung

Das Problem SAT ist das Erfüllbarkeitsproblem für Formeln der klassischen Aussagenlogik in konjunktiver Normalform (KNF).

Gegeben:	Eine aussagenlogische Formel F in KNF, in n Variablen.
Frage:	Ist F erfüllbar?

Bekanntlich ist dies das kanonische NP-vollständige Problem, daher gibt es keinen Algorithmus, der dieses effizient, also in polynomieller Zeit löst.

Auf der anderen Seite ist das Problem trivialerweise durch brute-force Ausprobieren aller 2^n möglichen Variablenbewertungen in exponentieller Zeit $O(2^n \cdot |F|)$ lösbar. Unter einer plausiblen Hypothese der Komplexitätstheorie, der *strong exponential time hypothesis* SETH, gibt es keinen Algorithmus, der es in subexponentieller Zeit $2^{o(n)} \cdot |F|$ löst.

Von einem Algorithms für SAT wird erwartet, dass er nicht nur das obige Entscheidungsproblem lößt, sondern im positiven Fall auch eine Bewertung α zurückgibt, die F erfüllt. Die Forschung über solche SAT Algorithmen hat sit der MItte der 1990er Jahre sehr viele Fortschritte gemacht, sowohl in der Theorie als aich in der Praxis.

Wegen der obigen komplexitätstheoretischen Einschränkungen ist es nicht zu erwarten, dass dabei Algorithmen mit einer echt subexponentiellen Laufzeit im worst case gefunden werden. Es gibt aber immer bessere Algorithmen, für die schwach exponentielle obere Schranken an die Laufzeit gezeigt werden können.

Die besten solchen Algorithmen sind probabilistisch, und haben eine Laufzeit von $O(1,307^n \cdot |F|)$ für 3-SAT und $O(1,469^n \cdot |F|)$ für 4-SAT. Die besten derzeit bekannten deterministischen Algorithmen erreichen $O(1,3303^n \cdot |F|)$ für 3-SAT und $O(1,5^n \cdot |F|)$ für 4-SAT.

Auf der anderen Seite gibt es seit Ende der 1990er Jahre andere Algorithmen, die in der Praxis sehr gut funktionieren, ihre Implementierungen werden als SAT Solver bezeichnet. Solche SAT Solver lösen Instanzen aus Anwendungen mit 100.000en von Variablen und Millionen von Klauseln in sehr kurzer Zeit. Sie sind mittlerweile so effinient, dass sie in vielen Bereichen als universelle Problemlöser verwendet werden.

Industriestandard ist dies bei der Verifikation von Hardware. Jeder neu entwickelte Chip wird heute bei allen Herstellern formal verifiziert, und die dabei verwendeten Methoden (Bounded Model Checking) verwenden SAT Solver als wesentlichen Bestandteil. Die Münchener Firma OneSpin Solutions GmbH, ursprünglich eine Ausgründung von Infineon, entwickelt diese Technologie und vermarktet sie an Chiphersteller.

Der Bereich der Erforschung von SAT Algorithmen ist mittlerweile ein eigener Forschungszweig mit einer wissenschaftlichen Gesellschaft die eine jährliche Konferenz veranstaltet und ein eigenes Journal herausgibt. Mit zum Erfolg des Gebietes tragen sicher auch die jährlich im Rahmen der Konferenz veranstalteten SAT Competitions bei, wo die Entwickler von SAT Solvern diese im Wettbewerb gegeneinander antreten lassen.

Wichtig für die Anwendung ist insbesondere, dass der Antwort eines SAT Solvers vertraut werden kann, also das Ergebnis unabhängig überprüfbar ist. Im positiven Fall ist dies einfach, da der Solver eine erfüllende Bewertung liefert, deren Korrektheit einfach durch Einsetzen in die Formel und Ausrechnen des Wertes geprüft werden kann.

Seit Beginn der 2010er Jahre wird von jedem SAT Solver verlangt, dass auch im negativen Fall ein Zerifikat, also ein Beweis für die Unerfüllbarkeit der Formel geliefert wird. Zu Beginn wurden dafür Resolutionsbeweise verwendet, ein SAT Solver hat also bei negativer Antwort einen Resolutionsbeweis für die Formel ausgegeben. Mittlerweile wurden effizientere Beweisformate entwickelt, für die es auch formal verifizierte Programme zur Überprüfung gibt.

Diese unabhängige Überprüfbarkeit erlaubt es, die von einem SAT Solver ausgegebenen Zertifikate für die Unerfüllbarkeit tatsächlich als Beweise im mathematischen Sinn aufzufassen. So wurden in jüngster Zeit einige neue mathematische Resultate mithilfe von SAT Solving Methoden bewiesen.

Prominentes Beispiel ist das Problem der Pythagoreischen Tripel. Ein Pythagoreisches Tripel ist bekanntlich ein Tripel von drei natürlichen Zahlen (a, b, c) mit $a^2 + b^2 = c^2$, also beispielsweise $(3, 4, 5)$ oder $(5, 12, 13)$. Eine Vermutung von Graham besagte, dass es für jede Färbung der natürlichen Zahlen mit zwei Farben ein einfarbiges pythagoreisches Tripel gibt. Diese Vermutung wurde 2016 von Heule, Kullmann und Marek mit SAT Solving

Methoden bewiesen.

Es lässt sich leicht eine CNF-Formel in Variablen $x_1, \dots x_N$ hinschreiben, die genau dann erfüllbar ist, wenn es eine Färbung der Zahlen $1, \dots, N$ gibt, die kein einfarbiges pythagoreisches Tripel enthält. Diese enthält für jedes solche Tripel (a, b, c) die Klauseln $x_a \vee x_b \vee x_c$ und $\neg x_a \vee \neg x_b \vee \neg x_c$.

Mit einem SAT Solver wurde gezeigt, dass diese Formel für $N \leq 7.824$ erfüllbar, für $N = 7.825$ jedoch unerfüllbar ist. Dafür brauchte ein speziell dafür getuneter paralleler SAT Solver ca. 2 Jahre an CPU Zeit. Der gelieferte Beweis der Unerfüllbarkeit ist etwa 2 TB groß und ging als längster mathematischer Beweis weltweit durch die Presse.

2 Grundlagen

2.1 Syntax der Aussagenlogik

Formeln der Aussagenlogik sind zusammengesetzt aus Variablen, von denen es eine abzählbar unendliche Menge X gibt, mittels der logischen Zeichen $\neg, \vee, \wedge$, sie sind induktiv definiert wie folgt:

1. Die Konstanten 0 und 1 sind Formeln.
2. Jede Variable $x \in X$ ist eine Formel.
3. Ist F eine Formel, so ist auch $\neg F$ eine Formel.
4. Sind F und G Formeln, so sind auch $F \vee G$ sowie $F \wedge G$ Formeln.

Variablen bezeichnen wir im Allgemeinen mit x, y, z und Formeln mit F, G, H, beide evtl. mit Indizes. Die Menge der in F vorkommenden Variablen wird mit $V(F)$ bezeichnet und kann induktiv definiert werden durch:

$$\begin{aligned} V(0) &= \emptyset & V(1) &= \emptyset \\ V(x) &= \{x\} & V(\neg F) &= V(F) \\ V(F \wedge G) &= V(F) \cup V(G) & V(F \vee G) &= V(F) \cup V(G) \end{aligned}$$

Die Größe $|F|$ einer Formel F ist die Anzahl der in F vorkommenden Zeichen, und kann induktiv definiert werden durch:

$$\begin{aligned} |0| &= 1 & |1| &= 1 \\ |x| &= 1 & |\neg F| &= |F| + 1 \\ |F \wedge G| &= |F| + |G| + 1 & |F \vee G| &= |F| + |G| + 1 \end{aligned}$$

Die Menge der $T(F)$ der Teilformeln einer Formel F ist induktiv definiert wie folgt:

$$\begin{aligned} T(0) &= \{0\} & T(1) &= \{1\} \\ T(x) &= \{x\} & T(\neg F) &= \{\neg F\} \cup T(F) \\ T(F \wedge G) &= \{F \wedge G\} \cup T(F) \cup T(G) & T(F \vee G) &= \{F \vee G\} \cup T(F) \cup T(G) \end{aligned}$$

2.2 Semantik der Aussagenlogik

Eine *Bewertung* ist eine endliche partielle Abbildung α, die Variablen $x \in X$ einen Wahrheitswert $\alpha(x) \in \{0, 1\}$ zuordnet. Wir beschreiben eine Bewertung durch eine Liste von Wertzuweisungen

$$[x_1 \leftarrow \epsilon_1, \ldots, x_k \leftarrow \epsilon_k]$$

für paarweise verschiedene Variable $x_1, \ldots, x_k$ und Werte $\epsilon_1, \ldots, \epsilon_k \in \{0, 1\}$. Bewertungen bezeichnen wir im Allgemeinen mit α, β, γ, evtl. mit Indizes.

Der Definitionsbereich von α wird mit $\operatorname{dom}\alpha$ bezeichnet. Eine Bewertung α heißt *total* für F, falls sie jede in F vorkommende Variable bewertet, also $V(F) \subseteq \operatorname{dom}\alpha$ ist, andernfalls heißt sie *partiell*.

Der Wert $F\alpha$ einer Formel F unter der Bewertung α wird errechnet, indem jede Variable $x \in \operatorname{dom}\alpha$ in F durch die Konstante $\alpha(x)$ ersetzt werden, und dann gemäß der folgenden Regeln vereinfacht wird:

$$\begin{array}{lll} \neg 0 \rightsquigarrow 1 & F \wedge 0\,,\ 0 \wedge F \rightsquigarrow 0 & F \vee 0\,,\ 0 \vee F \rightsquigarrow F \\ \neg 1 \rightsquigarrow 0 & F \wedge 1\,,\ 1 \wedge F \rightsquigarrow F & F \vee 1\,,\ 1 \vee F \rightsquigarrow 1 \end{array}$$

Falls α nicht total für F ist, ist $F\alpha$ eine Formel mit $V(F\alpha) \subseteq V(F) \setminus \operatorname{dom}\alpha$, kann aber auch ein Wert $F\alpha \in \{0, 1\}$ sein, z.B. $(x \vee y)[x \leftarrow 1] = 1$.

Eine Bewertung β *erweitert* α, in Zeichen $\beta \supseteq \alpha$, falls $\operatorname{dom}\alpha \subseteq \operatorname{dom}\beta$ ist, und für alle $x \in \operatorname{dom}\alpha$ gilt $\alpha(x) = \beta(x)$.

Falls für eine Formel F und Bewertung α gilt $F\alpha = 1$, so sagen wir auch α *erfüllt* F und schreiben dafür $\alpha \models F$.

Eine Formel F heißt

- *erfüllbar*, wenn es eine Bewertung α gibt mit $\alpha \models F$.
- *gültig* oder *Tautologie*, wenn $\alpha \models F$ für jede für F totale Bewertung α gilt.

Zwei Formeln F und G heißen *äquivalent*, in Zeichen $F \equiv G$, falls für jede Bewertung α, die für F und G total ist, gilt $F\alpha = G\alpha$. Dies läßt sich auch ausdrücken wie folgt: für jede totale Bewertung α gilt $\alpha \models F$ genau dann, wenn $\alpha \models G$.

Zwei Formeln F und G heißen erfüllbarkeits-äquivalent, in Zeichen $F \approx G$, falls gilt: F ist erfüllbar genau dann wenn G ist erfüllbar.

2.3 Normalformen

Ein *Literal* ist eine Variable x oder eine negierte Variable $\neg x$. Eine Variable x heißt auch *positives* Literal und eine negierte Variable $\neg x$ *negatives* Literal. Die Literale x und $\neg x$ sind *komplementär* zueinander. Literale bezeichnen wir im Allgemeinen mit $a, b, c, \ldots$, evtl. mit Indizes.

Die Notation für Bewertungen wird wie folgt auf Literale erweitert: Ist a die Variable x, so bedeutet $[a \leftarrow \epsilon]$ dasselbe wie $[x \leftarrow \epsilon]$, und ist a die negierte Variable $\neg x$, so bedeutet $[a \leftarrow \epsilon]$ dassselbe wie $[x \leftarrow (1 - \epsilon)]$.

Negations-Normalform

Eine Formel F ist in *Negations-Normalform* (NNF), falls Negationszeichen in F nur unmittelbar vor Variablen vorkommen, also falls F der folgenden induktiven Definition genügt:

- Jedes Literal und jede Konstante ist eine Formel in NNF.
- Sind F und G Formeln in NNF, so sind auch $F \vee G$ sowie $F \wedge G$ Formeln in NNF.

Für eine Formel F in NNF definieren wir die negierte Formel $\bar{F}$ in NNF durch:

- Ist $F = 0$, so ist $\bar{F} = 1$. Ist $F = 1$, so ist $\bar{F} = 0$.
- Ist $F = x$, so ist $\bar{F} = \neg x$.
- Ist $F = \neg x$, so ist $\bar{F} = x$.
- Ist $F = G_1 \vee G_2$, so ist $\bar{F} = \bar{G}_1 \wedge \bar{G}_2$.
- Ist $F = G_1 \wedge G_2$, so ist $\bar{F} = \bar{G}_1 \vee \bar{G}_2$.

Anschaulich entsteht $\bar{F}$ aus F dadurch, dass jedes Literal durch das entsprechende komplementäre Literal ersetzt wird, und die Symbole 0 und 1 sowie $\wedge$ und $\vee$ vertauscht werden.

Proposition 1. *Für jede Formel F in NNF ist $\bar{F}$ äquivalent zu $\neg F$.*

Beweis durch strukturelle Induktion nach der obigen induktiven Defintion der NNF.

- Ist $F = 0$, so ist $\bar{F} = 1 \equiv \neg 0$. Ist $F = 1$, so ist $\bar{F} = 0 \equiv \neg 1$.
- Ist $F = x$, so ist $\bar{F} = \neg x = \neg x$.
- Ist $F = \neg x$, so ist $\bar{F} = x \equiv \neg\neg x = \neg F$.
- Ist $F = G \vee H$, so ist $\bar{F} = \bar{G} \wedge \bar{H}$. Dies ist nach Induktionshypothese äquivalent zu $\neg G \wedge \neg H$, und nach dem De Morgan'schen Gesetz ist dies äquivalent zu $\neg(G \vee H) = \neg F$.
- Ist $F = G \wedge H$, so ist $\bar{F} = \bar{G} \vee \bar{H} \equiv \neg G \vee \neg H$ nach Induktionshypothese und dies ist äquivalent zu $\neg(G \wedge H) = \neg F$.

Proposition 2. *Für jede Formel F gibt es eine äquivalente Formel $n(F)$ in NNF.*

Beweis durch Induktion nach dem Formelaufbau:

- Ist $F = x$, $F = 0$ oder $F = 1$, so ist F bereits in NNF, also setze $n(F) = F$.
- Ist $F = G \circ H$ für $\circ \in \{\wedge, \vee\}$, so gibt es nach Induktionshypothese $n(G) \equiv G$ und $n(H) \equiv H$ in NNF, und wir setzen $n(F) = n(G) \circ n(H)$ und erhalten offensichtlich $n(F) = n(G) \circ n(H) \equiv G \circ H = F$.
- Ist $F = \neg G$ so gibt es nach Induktionshypothese $n(G) \equiv G$ in NNF. Definiere $n(F) = n(\bar{G})$, nach Prop. 1 gilt dann $n(F) = n(\bar{G}) \equiv \neg n(G) \equiv \neg G = F$.

Disjunktive und Konjunktive Normalform

Definition. *Wir definieren folgende spezielle Klassen von Formeln:*

1. *Ein* Term *ist eine Konjunktion* $a_1 \wedge \ldots \wedge a_k$ *von Literalen.*
2. *Eine* Klausel *ist eine Disjunktion* $a_1 \vee \ldots \vee a_k$ *von Literalen.*
3. *Eine Formel in* disjunktiver Normalform *(DNF) ist eine Disjunktion* $T_1 \vee \ldots \vee T_m$ *von Termen.*
4. *Eine Formel in* konjunktiver Normalform *(KNF) ist eine Konjunktion* $C_1 \wedge \ldots \wedge C_m$ *von Klauseln.*

Man beachte, dass jede Formel in DNF oder KNF auch in NNF ist.

Satz 3. *Für jede Formel* F *gibt es äquivalente Formeln* $\widehat{F}$ *in KNF und* $\check{F}$ *in DNF.*

Beweis. Für eine Variable x seien die Literale $x^1 = x$ und $x^0 = \neg x$ definiert. Für $\epsilon = 0, 1$ gilt also, dass $\alpha \models x^\epsilon$ genau dann wenn $\alpha(x) = \epsilon$ ist.
Wir definieren nun die Formel $\check{F}$. Sei $V(F) = \{x_1, \ldots, x_n\}$. Für eine Bewertung $\alpha = [x_1 \leftarrow \epsilon_1, \ldots, x_n \leftarrow \epsilon_n]$ betrachte den Term

$$T(\alpha) = x_1^{\epsilon_1} \wedge \ldots \wedge x_n^{\epsilon_n} .$$

Für jede Bewertung β gilt dann: $\beta \models T(\alpha)$ genau dann, wenn $\beta \supseteq \alpha$.
Es seien nun $\alpha_1, \ldots, \alpha_m$ alle Bewertungen mit $\operatorname{dom} \alpha_j = \{x_1, \ldots, x_n\}$ und $\alpha_j \models F$ für $j = 1, \ldots, m$. Wir definieren dann

$$\check{F} = T(\alpha_1) \vee \ldots \vee T(\alpha_m)$$

und müssen zeigen, dass $\check{F} \equiv F$ ist.

Sei also β eine für F totale Bewertung mit $\beta \models F$, dann gibt es ein $j = 1, \ldots, m$ so dass $\beta \supseteq \alpha_j$, also gilt $\beta \models T(\alpha_j)$, und somit $\beta \models \check{F}$.

Umgekehrt gelte $\beta \models \check{F}$, dann muss gelten $\beta \models T(\alpha_j)$ für ein $j = 1, \ldots, m$, und somit $\beta \supseteq \alpha_j$, also nach Definition $\beta \models F$.

Die Formel $\widehat{F}$ in KNF erhält man nun wie folgt: Bilde zunächst die zu $\neg F$ äquivalente Formel G in DNF, $G = \neg\check{F}$. Insbesondere ist G in NNF, also definieren wir $\widehat{F} = \bar{G}$, und nach Prop. 1 ist $\widehat{F} \equiv \neg G \equiv F$. □

Die im obigen Beweis konstruierte Formel $\check{F}$ hat die Größe $O(n2^n)$. Dass diese Konstruktion kann nicht wesentlich verbessert werden kann, zeigt die folgende untere Schranke:

Proposition 4. *Es gibt Formeln F_n in KNF mit $|F_n| = O(n)$, derart dass jede zu F_n äquivalente Formel G_n in DNF die Größe $|G_n| \geq n2^n$ hat.*

Beweis. Die Formeln F_n sind definiert durch

$$F_n = \bigwedge_{i=1}^{n} (x_{i,0} \vee x_{i,1}) \wedge (\neg x_{i,0} \vee \neg x_{i,1}) \; .$$

Eine Bewertung erfüllt die Formel F_n genau dann, wenn sie für jedes $i = 1, \ldots, n$ *genau eine* der Variablen $x_{i,0}$ und $x_{i,1}$ mit 1 bewertet.

Sei $G_n = T_1 \vee \ldots \vee T_m$ eine zu F_n äquivalente Formel in DNF. O.b.d.A. können wir annehmen, dass keiner der Terme T_j unerfüllbar ist, ansonsten könnte man diesen einfach weglassen und erhielte eine kleinere äquivalente Formel in DNF.

Wir zeigen zunächst, dass jeder Term in G_n für jedes $i = 1, \ldots, n$ genau eine der Variablen $x_{i,0}$ oder $x_{i,1}$ positiv enthalten muss.

Ist nämlich T_j ein Term in G_n, der weder $x_{i,0}$ noch $x_{i,1}$ positiv enthält, dann wähle eine Bewertung $\alpha \models T$ mit $\alpha(x_{i,0}) = \alpha(x_{i,1}) = 0$. Ist andererseits T_j ein Term, der sowohl $x_{i,0}$ als auch $x_{i,1}$ positiv enthält, dann wähle eine Bewertung $\alpha \models T$ mit $\alpha(x_{i,0}) = \alpha(x_{i,1}) = 1$. In beiden Fällen gilt $\alpha \models G_n$, aber $\alpha \not\models F_n$.

Ferner zeigen wir, dass für jedes $\epsilon = (\epsilon_1, \ldots, \epsilon_n) \in \{0,1\}^n$ ein Term in G_n existiert, der gerade die Variablen $x_{1,\epsilon_1}, \ldots, x_{n,\epsilon_n}$ positiv enthält. Ist dies nämlich für ein $\epsilon \in \{0,1\}^n$ nicht der Fall, dann muss jeder Term in G_n eine der Variablen $x_{1,(1-\epsilon_1)}, \ldots, x_{n,(1-\epsilon_n)}$ positiv enthalten. Dann wähle eine Bewertung α mit $\alpha(x_{i,\epsilon_i}) = 1$ und $\alpha(x_{i,(1-\epsilon_i)}) = 0$ für alle $i = 1, \ldots, n$. Für diese gilt $\alpha \models F_n$, aber $\alpha \not\models G_n$.

Die Formel G_n enthält also mindestens $m \geq 2^n$ Terme, von denen jeder mindestens n Literale hat, also ist $|G_n| \geq n2^n$. □

Korollar 5. *Es gibt Formeln F'_n in DNF mit $|F'_n| = O(n)$, derart dass jede zu F_n äquivalente Formel G_n in KNF die Größe $|G_n| \geq n2^n$ hat.*

Beweis. Sei $F'_n := \bar{F}_n$ für die Formeln F_n aus Proposition 4. Diese Formeln sind in DNF, und wenn es Formeln $G_n \equiv F'_n$ in KNF mit $G_n < n2^n$ gäbe, dann wäre auch $\bar{G}_n \equiv F_n$ eine Formel in DNF mit $|\bar{G}_n| = |G_n| < n2^n$, im Widerspruch zu Proposition 4. □

Die *Breite* $w(C)$ einer Klausel $C = a_1 \vee \ldots \vee a_k$ ist k, die Anzahl der Literale in C. Eine Klausel der Breite $w(C) = k$ wird auch als k-Klausel bezeichnet.

Definition. *Eine Formel $F = C_1 \wedge \ldots \wedge C_m$ in KNF ist in k-KNF für $k \in \mathbb{N}$, falls jede Klausel C_j die Breite $w(C_j) \leq k$ hat.*

Die k-KNF ist im Gegensatz zur vollen KNF keine echte Normalform, da nicht jede Formel in diese Form gebracht werden kann:

Proposition 6. *Für jedes k gibt es eine Formel F in $(k+1)$-KNF, für die es keine äquivalente Formel in k-KNF gibt.*

Beweis. Sei $F = x_1 \vee \ldots \vee x_{k+1}$. Angenommen, G sei eine zu F äquivalente Formel in k-KNF, also $G = C_1 \wedge \ldots \wedge C_m$. O.B.d.A. ist C_1 nicht tautologisch, und C_1 enthält höchstens k Variablen, also o.B.d.A. die Variable x_{k+1} nicht. Wähle also eine Bewertung α so, dass $\alpha(C_1) = 0$ ist, aber $\alpha(x_{k+1}) = 1$. Dann gilt $\alpha \models F$, aber $\alpha \not\models G$, im Widerspruch zur Annahme. □

Im folgenden werden vor allem Formeln in KNF für uns von Interesse sein. Für eine solche Formel F legen wir die folgenden Bezeichnungen fest:

n	die Anzahl der Variablen in F,
m	die Anzahl der Klauseln in F,
k	die maximale Breite einer Klausel in F.

Ferner identifizieren wir zur Vereinfachung der Notation Formeln in KNF mit Mengen von Klauseln. Eine Menge M von Klauseln stehe also für die Formel $F_M := \bigwedge_{C \in M} C$ und umgekehrt. Außerdem bedeute $F \setminus C$ für eine Klausel C dasselbe wie $F \setminus \{C\}$.

2.4 Das Problem SAT und seine Komplexität

Das Entscheidungsproblem FSAT ist definiert wie folgt:

Gegeben:	Eine Formel F der Aussagenlogik.
Frage:	Ist F erfüllbar?

Ohne Beweis zitieren wir das folgende zentrale Resultat:

Satz 7 (Cook 1971)**.** *FSAT ist NP-vollständig.*

Satz 8. *Für jede Formel* F *gibt es eine erfüllbarkeits-äquivalente Formel* $F' \approx F$ *in 3-KNF mit* $|F'| = O(|F|)$*. Weiter ist* F' *aus* F *in polynomieller Zeit berechenbar.*

Beweis. Wir definieren für jede Teilformel $G \in T(F)$ eine Variable x_G und eine Menge F_G von Klauseln wie folgt:

- Ist G eine Variable x, so ist $x_G = x$ und $F_G = \emptyset$.
- Ist $G = H \vee H'$, so ist x_G eine neue Variable, und F_G enthält Klauseln, die ausdrücken, dass $x_G \leftrightarrow (x_H \vee x_{H'})$ ist, nämlich $(\neg x_G \vee x_H \vee x_{H'})$, $(\neg x_H \vee x_G)$ und $(\neg x_{H'} \vee x_G)$).
- Ist $G = H \wedge H'$, so ist x_G eine neue Variable, und F_G enthält die Klauseln $(\neg x_H \vee \neg x_{H'} \vee x_G)$, $(\neg x_G \vee x_H)$ und $(\neg x_G \vee x_{H'})$).
- Ist $G = \neg H$, so ist x_G eine neue Variable, und F_G enthält die Klauseln $(x_G \vee x_H)$ und $(\neg x_G \vee \neg x_H)$.

Die Formel F' ist nun definiert als $\bigwedge_{G \in T(F)} F_G \wedge x_F$. Es bleibt zu zeigen, dass $F' \approx F$ ist.

Sei also $\alpha \models F'$. Durch Induktion nach dem Aufbau von F zeigen wir, dass für jede Teilformel $G \in T(F)$ gilt $\alpha(x_G) = G\alpha$. Insbesondere ist also $F\alpha = \alpha(x_F)$, und wegen des letzten Konjunktionsglieds muss $\alpha(x_F) = 1$ sein, somit gilt $\alpha \models F$.

Der Induktionsanfang $G = x$ ist trivial. Sei also z.B. $G = H \vee H'$. Nach Induktionshypothese ist $\alpha(x_H) = H\alpha$ und $\alpha(x_{H'}) = H'\alpha$. Die Klauseln F_G erzwingen nun, dass $\alpha(x_G) = \alpha(x_H) \vee \alpha(x_{H'})$ ist, also folgt $\alpha(x_G) = H\alpha \vee H'\alpha = G\alpha$. Die übrigen Fälle sind analog.

Sei umgekehrt $\alpha \models F$. Dann wird eine Bewertung $\alpha' \supseteq \alpha$ definiert durch $\alpha'(x_G) = G\alpha$ für jede Teilformel $G \in T(F)$. Dann rechnet man leicht nach, dass $\alpha' \models F_G$ für alle Formeln F_G gilt. Da $F\alpha = 1$ ist, gilt auch $\alpha'(x_F) = 1$, also ist insgesamt $F'\alpha' = 1$, also $\alpha' \models F'$. □

Das Entscheidungsproblem SAT ist der folgende Spezialfall von FSAT:

Gegeben:	Eine Formel F in KNF.
Frage:	Ist F erfüllbar?

Das Entscheidungsproblem k-SAT ist der folgende Spezialfall von SAT:

Gegeben: Eine Formel F in k-KNF.
Frage: Ist F erfüllbar?

Aus Satz 8 folgt unmittelbar, dass die Erfüllbarkeit von Formeln in 3-KNF ebenso schwierig ist wie für allgemeine Formeln.

Korollar 9. *3-SAT ist NP-vollständig. Daher sind auch SAT und* k*-SAT für jedes* $k \geq 3$ *NP-vollständig.*

Im Gegensatz dazu ist es für Formeln in DNF sehr einfach zu entscheiden, ob sie erfüllbar sind. Eine solche Formel F ist nämlich genau dann erfüllbar, wenn es einen Term in F gibt, der keine komplementären Literale a und $\bar{a}$ enthält. Daraus erhält man aber keinen effizienten Algorithmus für SAT, da die Umwandlung von Formeln in DNF aufwändig sein kann, wie Prop. 4 zeigt.

2.5 Beschränkte Vorkommen

Sei $\mathcal{F}$ eine Klasse von Formeln in KNF. Eine Formel $F \in \mathcal{F}$ ist in $\mathcal{F}(l)$ für $l \in \mathbb{N}$, wenn jede Variable in F höchstens l mal vorkommt. Das Problem SAT(l) ist das Erfüllbarkeitsproblem SAT für Formeln in KNF(l).

Satz 10. *Für jede Formel* F *in KNF gibt es eine erfüllbarkeits-äquivalente Formel* $D(F)$ *in KNF(3) mit* $w(D(F)) = w(F)$.

Beweis. Wir geben eine polynomielle Reduktion von SAT auf SAT(3) an. Sei F eine Formel in KNF. Für jede Variable x, die in F ℓ mal vorkommt, wählen wir ℓ neue Variablen $x_1, \ldots, x_\ell$, und ersetzen das i-te Vorkommen von x in F durch x_i, für $1 \leq i \leq \ell$. Das Ergebnis dieser Ersetzungen sei F'.

Damit F' zu F erfüllbarkeits-äquivalent wird, müssen wir erzwingen, dass die Variablen $x_1, \ldots, x_\ell$ immer gleich bewertet werden. Dazu fügen wir für jede Variable $x \in V(F)$ die Klauseln E_x hinzu:

$$E_x := (\bar{x}_1 \vee x_2) \wedge \ldots \wedge (\bar{x}_{\ell-1} \vee x_\ell) \wedge (\bar{x}_\ell \vee x_1)$$

Definiere nun die Formel $D(F) := F' \wedge \bigwedge_{x \in V(F)} E_x$. Offenbar ist $D(F)$ in KNF(3), denn jede Variable x_i kommt in $D(F)$ dreimal vor: einmal in F', und zweimal in E_x.

Jede Bewertung $\alpha \models F$ lässt sich nun zu einer $\alpha' \models D(F)$ umbauen, indem man für jede Variable x setzt $\alpha'(x_i) = \alpha(x)$ für $1 \leq i \leq \ell$.

Umgekehrt ist $\beta \models D(F)$, so müssen die Variabeln x_i für ein $x \in V(F)$ alle den gleichen Wert bekommen, da $\beta \models E_x$. Also definiere $\beta'(x) := \beta(x_1)$, dann folgt $\beta' \models F$, da ja $\beta \models F'$. □

Die Konstruktion ist eine polynomielle Reduktion, also folgt:

Korollar 11. *3-SAT(3) ist NP-vollständig.*

Die Reduktion im obigen Beweis erzeugt Klauseln der Breite 2. Die folgenden Überlegungen zeigen, dass dies notwendig ist.

Eine Formel in KNF ist in Ek-KNF, wenn jede Klausel genau die Breite k hat. Ek-SAT bezeichnet das Erfüllbarkeitsproblem für solche Formeln.

Satz 12. E3-*SAT ist NP-vollständig.*

Proof. Wir reduzieren 3-SAT auf E3-SAT. Sei $F = C_1 \wedge \ldots \wedge C_m$ in 3-KNF. Wir konstruieren daraus eine Formel F' in E3-KNF.

Ist C_i in F mit $w(C_i) = 2$, dann ersetzen wir C_i durch die zwei Klauseln $C_i \vee y_i$ und $C_i \vee \bar{y}_i$ der Breite 3 für eine neue Variable y_i. Analog werden Klauseln C_i der Breite 1 durch Verwendung zweier neuer Variablen durch vier Klauseln der Breite 3 ersetzt.

Wir zeigen, dass F'§ erfüllbar ist, genau dann wenn F erfüllbar ist. Offenbar gilt für jede Bewertung α mit $\alpha \models F$ auch $\alpha \models F'$.

Sei dagegen $\alpha \models F'$, und betrachte eine Klausel C_i in F der Breite 2. Ist $\alpha(y_i) = 1$, so muss $C_i\alpha = 1$ sein, da α auch $C_i \vee \bar{y}_i$ erfüllt. Analog gilt $\alpha \models C_i$ auch wenn $\alpha(y_i) = 0$. Für eine Klausel C_i in F der Breite 1 zeigt man $\alpha \models C_i$ auf die gleiche Weise, also gilt auch $\alpha \models F$. □

Analog zeigt man die NP-Vollständigkeit von Ek-SAT für jedes $k \geq 3$. Der folgende Satz zeigt, dass die Verwendung von Klauseln der Breite 2 in der Reduktion im Beweis von Satz 10 notwendig ist.

Satz 13. E3-*SAT(3) ist trivial, d.h., jede Formel in* E3-*KNF(3) ist erfüllbar.*

Sei $G = (U, V, E)$ ein bipartiter Graph mit $E \subseteq U \times V$. Eine Menge $M = \{(u_1, v_1), \ldots, (u_m, v_m)\}$ ist ein *Matching* in G, wenn die Kanten in M keine gemeinsamen Endpunkte haben, also wenn sowohl die u_i als auch die v_i paarweise verschieden sind. Ein Matching in G ist perfekt, wenn jeder Knoten in U mit einer Kante in M inzidiert, also wenn $\{u_1, \ldots, u_m\} = U$ ist. Ein perfektes Matching ordnet also jedem Knoten in U eindeutig einen Nachbarn in V zu.

Halls Theorem, auch bekannt als *Heiratssatz*, ist ein notwediges und hinreichendes Kriterium für die Existenz eines perfekten Matchings in G.

Für eine Teilmenge $S \subseteq U$ ist die Nachbarschaft $N(S) \subseteq V$ die Menge derjenigen $v \in V$, für die es eine Kante (u, v) zu einem $u \in S$ gibt.

Satz 14 (Halls Theorem). *Für einen bipartiten Graphen* $G = (U, V, E)$ *sind äquivalent:*

- *Es gibt ein perfektes Matching in* G*,*
- *Für jede Teilmenge* $S \subseteq U$ *ist* $|N(S)| \geq |S|$*.*

Beweis von Theorem 13. Für eine Formel $F = C_1 \wedge \ldots \wedge C_m$ ist der *Inzidenzgraph* der folgende bipartite Graph $G(F) = (U, V, E)$, wobei $U = \{C_1, \ldots, C_m\}$ die Klauseln in F sind, und $V = V(F)$ die Variablen in F sind. Zwischen Klausel C und einer Variablen x gibt es eine Kante $(C, x) \in E$, falls x in C vorkommt.

Sei nun $L \subseteq U$ eine Menge von Klauseln der Größe $s = |L|$, und $N(L) = V(L) \subseteq V$ die Menge ihrer Nachbarn, mit $t = |N(L)|$. Es bezeichne e die Anzahl der Kanten zwischen L und $N(L)$. Da F in E3-KNF ist, gilt $e = 3s$, und da F in KNF(3) ist, gilt $e \leq 3t$. Also ist $3t \geq 3s$ und somit $t \geq s$.

Nach Halls Theorem gibt es also ein perfektes Matching in $G(F)$, d.h., jeder Klausel C können wir eindeutig eine Variable $x_C \in V(C)$ zuordnen, deren Wert so gesetzt werden kann, dass C erfüllt ist. □

Auf der anderen Seite gilt:

Satz 15. E3-*SAT(4) ist NP-vollständig.*

Beweis. Wir reduzieren E3-SAT auf E3-SAT(4). Für jede Formel F in E3-KNF definieren wir eine Formel $D'(F)$ in E3-KNF(4), die genau dann erfüllbar ist, wenn F erfüllbar ist. Die Konstruktion basiert auf der im Beweis von Satz 10. Die Formel F' ist genau so definiert wie dort, aber statt der Formeln E_x, die die Äquivalenz der Kopien $x_1, \ldots, x_\ell$ einer Variable x erzwingen und aus Klauseln der Breite 2 bestehen, konstruieren wir Formeln E'_x mit dem selben Zweck.

Für jede Variable x werden ℓ neue Variablen $y_1, \ldots, y_\ell$ eingeführt, und die Klauseln in E_x ersetzt durch

$$E'_x := (\bar{x}_1 \vee x_2 \vee y_1) \wedge \ldots \wedge (\bar{x}_{\ell-1} \vee x_\ell \vee y_{\ell-1}) \wedge (\bar{x}_\ell \vee x_1 \vee y_\ell)$$

Für jede Variable x und jedes $i \leq \ell$ werden nun weitere Klauseln hinzugefügt, die erzwingen, dass die Variable y_i in jeder erfüllenden Bewertung mit 0 bewertet wird. Dann erfüllen die Klauseln E'_x denselben Zweck wie E_x im Beweis von Satz 10.

Dafür werden für jede Variable x und jedes $i \leq \ell$ noch vier neue Variablen a_i, b_i, c_i, d_i eingeführt, und dazu die sechs Klauseln

$$\bar{y}_i \vee a_i \vee \bar{b}_i \qquad a_i \vee b_i \vee \bar{c}_i \qquad b_i \vee c_i \vee d_i$$

$$b_i \vee c_i \vee \bar{d}_i \qquad \bar{y}_i \vee \bar{a}_i \vee d_i \qquad \bar{y}_i \vee \bar{a}_i \vee \bar{d}_i$$

Man überprüft leicht, dass diese sechs Klauseln unerfüllbar sind, wenn y_i mit 1 bewertet wird. Ausserdem hat jede Klausel die Breite 3, und es kommt jede Variable höchstens 4 mal vor, also ist die konstruierte Formel $D'(F)$ in E3-KNF(4).
Offenbar ist $D'(F)$ aus F in polynomieller Zeit konstruierbar, also ist dies eine polynomielle Reduktion von E3-SAT auf E3-SAT(4). □

Der Beweis von Theorem 13 ist leicht für beliebige Breiten k zu verallgemeinern: für jedes $k \geq 2$ ist Ek-SAT(k) trivial. Dagegen verallgemeinert sich Theorem 15 nicht direkt, Ek-SAT$(k+1)$ ist nicht für jedes k schwierig. Allerdings gilt, dass der abrupte Sprung von trivial zu NP-schwer für jedes k bei einer gewissen Anzahl s von Vorkommen auftritt:

Satz 16. *Für jedes* $k \geq 3$ *gibt es ein* $s = s(k) \geq k$ *mit:*

- Ek-*SAT(s) ist trivial,*
- Ek-*SAT(*$s+1$*) ist NP-vollständig.*

Der genaue Wert von $s(k)$ ist nur für kleine Werte von k bekannt. Für hinreichend große k gelten die Abschätzungen $2^k/ek \leq s(k) < 4 \cdot 2^k/k$. Die besten bekannten konkreten unteren und oberen Schranken für kleine Werte von k sind in der folgenden Tabelle zusammengefasst:

k	$s(k) \geq$	$s(k) <$
3	3	4
4	4	5
5	5	8
6	7	12
7	13	18
8	24	30
9	41	52

Wegen des abrupten Sprunges von trivial zu NP-schwer reicht es z.B. zur Verbesserung der oberen Schranke in der dritten Zeile, eine einzige unerfüllbare Formel in (E5)-KNF mit höchstens 7 Vorkommen jeder Variablen zu finden.

2.6 Resolution

Ein Resolutionsherleitung einer Klausel C aus einer Formel F in KNF ist eine Folge $C_1, \ldots, C_s$ von Klauseln mit den Eigenschaften:

- $C_s = C$
- für alle $1 \leq i \leq s$ gilt eine der folgenden Aussagen:
  - C_i ist eine Klausel in F
  - es gibt $j < j' < i$, so dass C_i Resolvente von C_j und $C_{j'}$ ist.

Ein Resolutionsbeweis für F ist eine Herleitung der leeren Klausel 0 aus F.

Satz 17. *Eine Formel* F *hat einen Resolutionsbeweis genau dann, wenn* F *unerfüllbar ist.*

Beweis. Im Kapitel 4 wurde folgende Eigenschaft gezeigt: Ist F erfüllbar, und C mittels Resolution aus F herleitbar, dann ist auch $F \wedge C$ erfüllbar. Ist also die unerfüllbare leere Klausel aus F herleitbar, so muss F unerfüllbar sein, was die Korrektheit zeigt.
Zur Vollständigkeit wird durch Induktion nach der Anzahl $|V(F)|$ der Variablen in F bewiesen, dss F einen Resolutionsbeweis hat, falls F unerfüllbar ist.
Hat F nur eine Variable x, so müssen die Einerklauseln x und $\bar{x}$ in F vorkommen, wenn F unerfüllbar ist. Dann folgt mittels eines Resolutionsschrittes aus diesen die leere Klausel, somit gibt es einen Resolutionsbeweis.
Sei also $|V(F)| = n + 1$, und gelte die Aussage für Formeln mit höchstens n Variablen. Wähle eine Variable x, so dass sowohl x als auch $\bar{x}$ in F vorkommen. Gibt es eine solche Variable nicht, muss F erfüllbar sein.
Betrachte die Formeln $F_0 := F[x := 0]$ und $F_1 := F[x := 1]$. Diese Formeln sind beide unerfüllbar, und sie haben jeweils höchstens n Variablen. Nach Induktionshypothese haben also beide Resolutionsbeweise.
Sei also $C_1, \ldots, C_s$ ein Resolutionsbeweis für F_0. Wir ersetzen jede Klausel C_i durch eine Klausel $C'_i \subseteq C_i \vee x$ wie folgt:

- Ist $C_i \in F$, dann ist $C'_i = C_i$.
- Ist C_i derart dass $C_i \vee x$ in F vorkommt, dann setze $C'_i = C_i \vee x$.
- Andernfalls ist C_i Resolvente von C_{j_1} und C_{j_2} für $j_1, j_2 < i$.
  - Ist $C'_{j_1} = C_{j_1}$ und $C'_{j_2} = C_{j_2}$, dann setze $C'_i = C_i$.
  - Ist dagegen $C'_j = C_j \vee x$ für ein $j = j_1, j_2$, dann setze auch $C'_i = C_i \vee x$.

Jede Klausel C'_i ist entweder in F, oder eine Resolvente von C'_{j_1} und C'_{j_2}, für $j_1, j_2 < i$. Also ist die Folge eine Resolutionsherleitung. Die letzte Klausel

C'_s ist entweder leer, dann liegt ein Resolutionsbeweis von F vor, oder die Einerklausel x.

Im zweiten Fall wird analog aus einem Resolutionsbeweis von F_1 eine Herleitung der Einerklausel $\bar{x}$ konstruiert. Fügt man nun diese beiden Herleitungen hintereinander, und fügt die leere Klausel als Resolvente von x und $\bar{x}$ hinzu, erhält man einen Resolutionsbeweis für F. □

3 Einfache Spezialfälle

Wir betrachten einige Spezialfälle von Formeln, für die es effiziente Algorithmen für das Erfüllbarkeitsproblem gibt. Eine Klasse $\mathcal{F}$ von Formeln ist ein einfacher Spezialfall von SAT, wenn die folgenden Bedingungen gelten:

- Es ist in polynomieller Zeit entscheidbar, ob eine Formel F in $\mathcal{F}$ liegt.
- Für Formeln $F \in \mathcal{F}$ ist die Erfüllbarkeit in polynomieller Zeit entscheidbar.

Wir betrachten im Folgenden einige klassische einfache Spezialfälle.

Horn-Formeln

Definition. *Eine Klausel heißt*

- positiv *(*negativ*), wenn sie nur positive (negative) Literale enthält,*
- Horn-Klausel, *wenn sie höchstens ein positives Literal enthält.*
- definit, *wenn sie genau ein positives Literal enthält.*

Jede Horn-Klausel ist also entweder definit oder negativ.
Eine Horn-Formel *ist eine Konjunktion von Horn-Klauseln, also eine Formel in KNF, in der jede Klausel eine Horn-Klausel ist.*

Horn-SAT ist der folgende Spezialfall von SAT:

Gegeben:	Eine Horn-Formel F.
Frage:	Ist F erfüllbar?

Jede unerfüllbare Formel enthält mindestens eine positive Klausel, denn sonst enthält jede Klausel mindestens ein negatives Literal, somit wird die Formel von der Bewertung erfüllt, die alle Variablen mit 0 bewertet. Analog gilt, dass jede unerfüllbere Formel auch eine negative Klausel enthält.
Daher muss jede unerfüllbere Horn-Formel eine positive 1-Klausel enthalten. Auf dieser Beobachtung basiert der unten angegebene Algorithmus für Horn-SAT.

Satz 18. *Horn-SAT ist polynomialer Zeit* $O(nm)$ *entscheidbar.*

Beweis. Der folgende Algorithmus entscheidet, ob eine eingegebene Horn-Formel F erfüllbar ist, und gibt gegebenenfalls eine erfüllende Bewertung aus.

Horn-Sat(F)
  $\alpha := []$
  while Fα enthält eine positive 1-Klausel
    wähle eine solche x
    $\alpha := \alpha \cup [x \leftarrow 1]$
  if $\alpha' \models F$
    then return α'
    else return unerfüllbar

wobei α' definiert ist durch

$$\alpha'(x) = \begin{cases} 1 & \alpha(x) = 1 \\ 0 & x \notin \operatorname{dom} \alpha \end{cases} .$$

Wir zeigen zuerst, dass für jede definite Klausel $C = x_1 \vee \bar{x}_2 \vee \ldots \vee \bar{x}_k$ in F gilt $\alpha' \models C$. Ist $\alpha'(x_i) = 0$ für ein $i = 2, \ldots, k$, dann gilt $\alpha' \models C$. Andernfalls kommt zu dem Zeitpunkt, wenn alle $x_2, \ldots, x_k$ in $\operatorname{dom} \alpha$ sind, die 1-Klausel x_1 in Fα vor, und somit setzt der Algorithmus $\alpha(x_1) = 1$. Dann gilt aber auch $\alpha' \models C$. Es bezeichne nun F_D die Menge der definiten Klauseln in F.

Wir zeigen nun, dass jede Bewertung $\beta \models F_D$ eine Erweiterung $\beta \supseteq \alpha$ ist. Andernfalls sei x die erste Variable, für die der Algorithmus $\alpha(x) = 1$ setzt, und für die $\beta(x) = 0$ ist. Zum Zeitpunkt, wo $\alpha(x) = 1$ gesetzt wird, kommt die 1-Klausel x in Fα vor, also gibt es eine definite Klausel $C = x \vee \bar{y}_1 \vee \ldots \vee \bar{y}_k$ in F, so dass die Variablen $y_1, \ldots, y_k$ bereits in $\operatorname{dom} \alpha$ sind. Nach Wahl von x ist $\beta(y_1) = \ldots = \beta(y_k) = 1$, also ist $\beta \not\models C$.

Schließlich zeigen wir, dass F unerfüllbar ist, falls $\alpha' \not\models F$, woraus sofort die Korrektheit des Algorithmus folgt. Falls $\alpha' \not\models F$, gibt es eine negative Klausel $C = \bar{y}_1 \vee \ldots \vee \bar{y}_k$ in F mit $\alpha' \not\models C$. Also sind $y_1, \ldots, y_k \in \operatorname{dom} \alpha$, und daher ist für jede Bewertung $\beta \models F_D$ auch $\beta(C) = 0$. Damit ist F unerfüllbar.

Offenbar ist die Zahl der Schleifendurchläufe durch n beschränkt, und in jedem Schleifendurchlauf wird nach einer positiven 1-Klausel gesucht. Durch Verwendung geeigneter Datenstrukturen (s.u.) ist dies in Zeit $O(m)$ möglich, der gesamte Zeitaufwand ist also $O(nm)$. □

Insbesondere liegt also Horn-SAT in der Klasse P, und ist sogar vollständig für diese Klasse.

2-SAT

Im Kontrast zu Korollar 9 ist das Problem 2-SAT einfach. Für die Konstruktion benötigen wir einige Begriffe und Algorithmen für gerichtete Graphen.

Sei $G = (V, E)$ ein gerichteter Graph mit $E \subseteq V \times V$. Ein Knoten v ist von u erreichbar, wenn es einen gerichteten Weg von u nach v gibt, also eine Folge von Knoten $v_1, \ldots, v_\ell$ mit $u = v_1$ und $v = v_\ell$ und $(v_i, v_{i+1}) \in E$ für $1 \leq i < \ell$.

Zwischen zwei Knoten $u, v \in V$ wird die Relation $u \sim v$ definiert durch: v ist von u erreichbar, und u ist von v erreichbar, d.h., u und v liegen auf einem Kreis. Diese Relation ist offenbar eine Äquivalenzrelation, die Äquivalenzklassen dieser Relation heißen *starke Zusammenhangskomponenten*. Die Äquivalenzklasse, die den Knoten $v \in V$ enthält, wird mit $[v]$ bezeichnet.

Es gibt einen Algorithmus, der in linearer Zeit (in der Anzahl der Knoten und Kanten) einen gerichteten Graphen in seine starken Zusammenhangskomponenten zerlegt.

Der Faktorgraph $G/\sim$ ist der Graph, dessen Knoten die Zusammenhangskomponenten $\{[v]; v \in V\}$ von G sind, mit einer Kante zwischen $[u]$ und $[v]$, wenn es $u' \in [u]$ und $v' \in [v]$ gibt mit $(u', v') \in E$. Dieser Graph ist also ein gerichteter, azyklischer Graph.

Für einen solchen gerichten, azyklischen Graphen gibt es eine *topologische Ordnung*, d.h. eine Anordnung $v_1, \ldots, v_n$ der Knoten so dass alle Kanten in der Ordnung nach rechts gehen: ist also $(v_i, v_j) \in E$, so ist $i < j$. Eine solche topologische Ordnung kann in linearer Zeit berechnet werden.

Satz 19. *2-SAT ist in linearer Zeit* $O(n + m)$ *entscheidbar.*

Beweis. Für jede Formel F in 2-KNF wird ein gerichteter Graph $G(F)$ definiert wie folgt:

- Die $2n$ Knoten von $G(F)$ sind die Literale x und $\neg x$ für jede Variable $x \in V(F)$.
- In $G(F)$ gibt es eine Kante $a \to b$, wenn die Klausel $\bar{a} \vee b$ in F vorhanden ist.
- Außerdem gibt es die Kanten $\bar{a} \to a$ für jede 1-Klausel a.

Jede Zweierklausel $a \vee b$ entspricht also zwei Kanten $\bar{a} \to b$ und $\bar{b} \to a$.

Wir zeigen die folgende Behauptung:

> Ist b von a in $G(F)$ erreichbar, und ist $\alpha \models F$ mit $\alpha(a) = 1$, dann ist auch $\alpha(b) = 1$.

Es genügt, dies für den Fall zu zeigen, wo eine Kante $a \to b$ vorhanden ist, die allgemeine Behauptung folgt induktiv. Für jede solche Kante ist aber

eine Klausel $\bar{a} \vee b$ in F vorhanden, die nicht erfüllt ist, wenn $\alpha(a) = 1$ und $\alpha(b) = 0$ ist.

Bezeichne nun $[a]$ für jedes Literal a die starke Zusammenhangskomponente von $G(F)$, in der der Knoten a liegt. Es gilt also: ist $\alpha \models F$, und $[a] = [b]$, so ist $\alpha(a) = \alpha(b)$. Daher ist F unerfüllbar, falls $[x] = [\bar{x}]$ für eine Variable x ist. Umgekehrt gilt: ist $[x] \neq [\bar{x}]$ für jede Variable x, dann ist F erfüllbar. Eine Bewertung $\alpha \models F$ wird wie folgt konstruiert:

Seien $[a_1], \ldots, [a_r]$ die starken Zusammenhangskomponenten in umgekehrter topologischer Ordnung (des induzierten gerichteten azyklischen Graphen).

for $j := 1$ to r do
    falls die Literale in $[a_j]$ noch unbewertet sind
        $\alpha(b) := 1$ für $b \in [a_j]$
        $\alpha(b) := 0$ für $b \in [\bar{a}_j]$

Nach der Voraussetzung ist diese Bewertung wohldefiniert. Offenbar erfüllt α alle Klauseln, die Kanten innerhalb einer Komponente entsprechen.

Sei $a \to b$ also eine Kante zwischen zwei Komponenten, und nehmen wir $\alpha \not\models \bar{a} \vee b$ an. Dann müssen die Literale in $[a]$ mit 1 und die in $[b]$ mit 0 bewertet worden sein. Deshalb liegt in der topologischen Ordnung $[\bar{a}]$ vor $[a]$ und $[\bar{b}]$ hinter $[b]$, und daher auch $[\bar{b}]$ hinter $[\bar{a}]$. Dies ist ein Widerspruch, da $G(F)$ auch die Kante $\bar{b} \to \bar{a}$ enthält.

Da die starken Zusammenhangskomponenten eines Graphen in linearer Zeit $O(n + m)$ bestimmt und topologisch angeordnet werden können, ist diese Konstruktion und der Test auf Erfüllbarkeit in linearer Zeit $O(m)$ möglich. □

Der Algorithmus zeigt sogar, dass das Problem 2-SAT in der Komplexitätsklasse NL liegt, also von einem nichtdeterministischen Algorithmus mit logarithmischem Speicherplatzbedarf gelöst werden kann. Tatsächlich ist 2-SAT auch vollständig für NL (unter einem geeigneten schwachen Reduktionsbegriff).

SAT(2)

Analog zum vorherigem Abschnitt steht das folgende Ergebnis im Kontrast zu Korollar 11:

Satz 20. *SAT(2) ist in linearer Zeit* $O(n)$ *entscheidbar.*

Beweis. Der folgende sehr einfache Algorithmus entscheidet SAT(2):

while in F gibt es eine 1-Klausel a
    $F := F[a \leftarrow 1]$
if $F = 0$
    then return UNSAT
    else return SAT

Nach Verlassen der Schleife enthält F keine 1-Klausel mehr. Enthält F nun eine leere Klausel, ist also $F = 0$, so ist sie offensichtlich unerfüllbar. Andernfalls hat jede Klausel C in F die Breite $w(C) \geq 2$, und jede Variable kommt höchstens zweimal vor. Daher ist F erfüllbar, wie dasselbe Matching-Argument wie für die Erfüllbarkeit von Formeln in E3-KNF(3), Satz 13, zeigt.

Ein anderer Ansatz zur Entscheidung von SAT(2) ist in gewissem Sinne dual zu dem Algorithmus für 2-SAT.

Ein *markierter Graph* $G = (V, E, M)$ ist ein ungerichteter Multigraph (V, E) (es kann also mehrere Kanten zwischen zwei Knoten geben) mit einer ausgezeichneten Teilmenge $M \subseteq V$ vom *markierten* Knoten. Eine Zusammenhangskomponente von G heiße *markiert*, falls sie mindestens einen markierten Knoten enthält.

Für eine Formel F in KNF(2) definiere einen markierten Graphen $G(F)$ wie folgt:

- $G(F)$ hat m Knoten, einen v_C für jede Klausel C in F.
- Falls die Klauseln C und D komplementäre Literale x und $\bar{x}$ enthalten, so gibt es eine Kante e_x zwischen v_C und v_D.
- Enthält C ein Literal a, derart dass $\bar{a}$ in F nicht vorkommt (ein *reines* Literal), so ist v_C markiert.

Man beachte, dass es zwischen v_C und v_D mehrere Kanten geben kann, falls C und D mehrere Paare komplementärer Literale enthalten, z.B. $C = (x \vee \bar{y})$ und $D = (\bar{x} \vee y \vee z)$.

Eine Bewertung der Variablen in F kann man nun so auffassen, dass sie den Kanten in $G(F)$ eine Richtung gibt: die Kante e_x wird so gerichtet, dass sie von derjenigen Klausel, die das mit 1 bewertete Literal aus $x, \bar{x}$ enthält, zu der Klausel, die das mit 0 bewertete Literal enthält, zeigt. Eine Klausel C ist dann erfüllt, wenn in dieser Ausrichtung v_C eine ausgehende Kante hat. Da Klauseln mit reinen Literalen stets erfüllt werden können, gilt die folgende offensichtliche Behauptung:

Proposition 21. *Eine Formel F in KNF(2) ist genau dann erfüllbar, genau dann wenn die Kanten in $G(F)$ so gerichtet werden können, dass es keine unmarkierte Senke gibt.*

Dies führt uns zum folgenden Lemma:

Lemma 22. *Eine Formel* F *in KNF(2) ist genau dann erfüllbar, wenn jede Zusammenhangskomponente in* $G(F)$ *markiert ist oder einen Kreis enthält.*

Beweis. Es reicht zu zeigen, dass die Bedingung des Lemmas äquivalent zu der Bedingung in Proposition 21 ist. Sie ist offensichtlich notwendig, da in einem endlichen gerichteten Graphen jeder Weg in einer Senke oder einem Kreis enden muss.

Ferner ist die Bedingung auch hinreichend: wir konstruieren eine passende Ausrichtung der Kanten für jede Zusammenhangskomponente, die entweder markiert ist oder einen Kreis enthält.

In einer markierten Komponente wird eine Ausrichtung wie folgt vorgenommen:

- Betrachte einen Spannbaum T der Komponente.
- Jede Kante in T wird so gerichtet, dass sie von dem von v weiter entfernten zum näher an v liegenden Knoten führt.

Die übrigen Kanten können dann beliebig gerichtet werden, da jeder Knoten außer der Wurzel v bereits eine ausgehende Kante in T hat.

In einer Komponente K , die einen Kreis $v_1, \ldots, v_k, v_{k+1} = v_1$ enthält, wird eine Ausrichtung wie folgt vorgenommen:

- Zuerst werden die Kanten in diesem Kreis um diesen herum gerichtet, also von v_1 zu v_2, ..., von v_{k-1} zu v_k und von v_k zu v_1.
- Sei E' die Menge der Kanten in K ohne diejenigen, die die Knoten $v_1, \ldots, v_k$ verbinden (also ohne den Kreis und seine Sehnen).
- Für jedes $1 \leq i \leq k$ sei K_i die Menge der Knoten, die von v_i, aber keinem v_j für $j < i$ aus in E' erreichbar sind.
- Für jedes i sei T_i ein Spannbaum des induzierten Graphen auf K_i. Die Kanten in T_i werden wie oben auf die Wurzel v_i hin gerichtet wie oben.

Auf diese Weise erhält jeder Knoten in K eine ausgehende Kante. Alle übrigen Kanten können dann beliebig gerichtet werden. □

Die Bedingung in Lemma 22 läßt sich leicht mittels Tiefensuche überprüfen, etwa mit dem folgenden Algorithmus:

solange noch unarkierte Knoten in G
    wähle den kleinsten unmarkierten Knoten v
    führe Tiefensuche ab v durch
    falls keinen markierten Knoten und keine rückläufige Kante gefunden
        return FALSE
return TRUE

Daher kann man die Erfüllbarkeit in Zeit $O(n+m)$ testen, dies ist von der Größenordnung $O(n)$, da $m \leq 2n$ ist. □

Tatsächlich liegt das Problem SAT(2) sogar in der Komplexitätsklasse L, kann also von einem deterministischen Algorithmus mit logarithmischem Speicherplatzbedarf gelöst werden, und ist auch vollständig für L (wiederum unter einem geeigneten schwachen Reduktionsbegriff).

Horn-umbenennbare Formeln

Eine *Umbenennung* von F ist eine Permutation r der Menge der Literale von F mit $r(a) \in \{a, \bar{a}\}$ für jedes Literal a. Eine Umbenennung ändert also nur die Vorzeichen mancher Literale.

Eine Formel F ist Horn-umbenennbar, wenn es eine Umbenennung r von F gibt, so dass $r(F)$ eine Horn-Formel ist.

Satz 23. *Es gibt einen Algorithmus, der in Zeit* $O(n^2m)$ *entscheidet, ob eine Formel* F *Horn-umbenennbar ist und ggf. ein Umbenennung* r *berechnet, für die* $r(F)$ *eine Horn-Formel ist.*

Beweis. Definiere für F die folgende Formel F^* in 2-KNF:

$$F^* := \{\, a \vee b \;;\; \text{es gibt } C \in F \text{ mit } a \in C \text{ und } b \in C \,\}$$

Es gilt, dass F^* genau dann erüllbar ist, wenn F Horn-umbenennbar ist, und aus einer Bewertung, die F^* erfüllt, gewinnt man eine Umbenennung die dies bezeugt.

Sei also $\alpha \models F^*$, daraus definieren wir die Umbenennung r_α mit

$$r_\alpha(x) = \begin{cases} \bar{x} & \text{falls } \alpha(x) = 1 \\ x & \text{sonst.} \end{cases}$$

Nun gilt: Ist $r_\alpha(a)$ positiv, so ist $\alpha(a) = 0$. Denn ist a positiv, so ist $r_\alpha(a) = a$, also $\alpha(a) = 0$. Ist dagegen $a = \bar{x}$, so ist $r_\alpha(a) = x = \bar{a}$, also ist $\alpha(x) = 1$ und somit $\alpha(a) = 0$.

Es folgt dass $r_\alpha(F)$ eine Horn-Formel ist. Denn andernfalls sei $C = a_1 \vee \ldots \vee a_k$ eine Klausel, für die $r_\alpha(C)$ keine Horn-Klausel ist. Dann gibt es Literale a_i und a_j in C, so dass $r_\alpha(a_i)$ und $r_\alpha(a_j)$ positiv sind, also ist $\alpha(a_i) = \alpha(a_j) = 0$, und somit ist die Klausel $a_i \vee a_j$ in F^* von α nicht erfüllt, im Widerspruch zur Voraussetzung.

Sei umgekehrt r eine Umbenennung, für die $r(F)$ eine Horn-Formel ist. Daraus definieren wir die Bewertung α_r mit

$$\alpha_r(x) = \begin{cases} 1 & \text{falls } r(x) = \bar{x} \\ 0 & \text{sonst.} \end{cases}$$

Nun gilt für jedes Literal a: ist $\alpha_r(a) = 0$, so ist $r(a)$ positiv. Denn ist $a = x$, so ist $\alpha_r(a) = \alpha_r(x) = 0$, also ist $r(x) = x$. Ist dagegen $a = \bar{x}$, so ist $\alpha(x) = 1$, also ist $r(a) = \bar{a} = x$.

Es folgt dass $\alpha_r \models F^*$. Denn andernfalls sei $a \vee b$ eine unerfüllte Klausel in F^*, dann ist $\alpha_r(a) = \alpha_r(b) = 0$. Dann gibt es aber eine Klausel $C \in F$ mit $(a \vee b) \subseteq C$, und $r(a)$ und $r(b)$ sind beide positiv, also ist $r(C)$ keine Horn-Klausel, im Widerspruch zur Voraussetzung.

Der Algorithmus konstruiert nun aus F die Formel F^*, die hat die Größe $O(n^2m)$. Dann löst er diese mit dem Linearzeit-Algorithmus für 2-SAT und gewinnt im Falle der Erfüllbarkeit die Umbenennung nach der obigen Definition. □

Ein Algorithmus für Horn-umbenennbare Formeln folgt daraus wie folgt:

- Berechne eine Umbenennung r so dass $r(F)$ eine Horn-Formel ist.
- Löse $r(F)$ mit dem Algorithmus für Horn-Formeln.
- Berechne ggf. aus der erfüllenden Bewertung für $r(F)$ eine für F durch Umkehrung der Umbenennung.

Die Laufzeit ist die zur Berechnung der Umbenennung $O(n^2m)$, da dies der aufwändigste Schritt ist.

Cluster-Formeln

Eine Formel F in KNF ist eine *Hitting*-Formel, wenn je zwei Klauseln in F komplementäre Literale enthalten, d.h., für je zwei $C, D \in F$ gibt es ein Literal a so dass $a \in C$ und $\bar{a} \in D$ ist.

Eine *Cluster*-Formel ist eine Konjunktion $F = H_1 \wedge \ldots \wedge H_\ell$ von Hitting-Formeln mit disjunkten Variablen, d.h. für $i \neq j$ ist $V(H_i) \cap V(H_j) = \emptyset$.

Die Erfüllbarkeit von Hitting-Formeln lässt sich leicht arithmetisch testen:

Satz 24. *Sei* $F = C_1 \wedge \ldots \wedge C_m$ *eine Hitting-Formel. Dann ist* F *genau dann erfüllbar, wenn gilt:*

$$\sum_{i=1}^{m} 2^{-w(C_i)} < 1 \,.$$

Beweis. Für eine Klausel $C_i \in F$ sei A_i die Menge der totalen Bewertungen α mit $C_i\alpha = 0$.

Behauptung: Für $i \neq j$ ist $A_i \cap A_j = \emptyset$.

Sei $\alpha \in A_i$. Ist $C_i\alpha = 0$, so ist $\alpha(a) = 0$ für alle $a \in C_i$. Nun ist aber $\bar{a} \in C_j$ für ein $a \in C_i$, also ist $C_j\alpha = 1$ und somit $\alpha \notin C_j$.

Sei ferner A die Menge der totalen Bewertungen α mit $\alpha \not\models F$. Also ist F genau dann erfüllbar, wenn $|A| < 2^n$.

Wegen der Disjunktheit ist $|A| = \sum_{i=1}^{m} |A_i|$, also ist F genau dann erfüllbar, wenn $\sum_{i=1}^{m} |A_i| < 2^n$ ist. Da $|A_i| = 2^{n-w(C_i)}$ ist, ist dies genau dann der Fall, wenn $\sum_{i=1}^{m} 2^{n-w(C_i)} < 2^n$ ist, also wenn $\sum_{i=1}^{m} 2^{-w(C_i)} < 1$ ist. □

Da eine Cluster-Formel $F = H_1 \wedge \ldots \wedge H_\ell$ erfüllbar ist, wenn alle H_i erfüllbar sind, ist diese arithmetische Bedingung für alle H_i zu überprüfen.

4 DPLL-Algorithmen

Offensichtlich ist SAT trivialerweise in Zeit $2^n \cdot O(|F|)$ durch Probieren aller 2^n möglichen Bewertungen lösbar. Dieses Ausprobieren läßt sich am besten durch ein Backtracking-Verfahren, in dem die vorkommenden Variablen sukzessive bewertet werden, organisieren. Solche Backtracking-Verfahren und ihre Verfeinerungen nennen wir DPLL-Algorithmen, nach den Autoren Davis, Putnam, Logemann und Loveland, die das erste Verfahren dieser Art 1960-1962 beschrieben haben.

Die grundlegende Struktur eines DPLL-Algorithmus kann wie folgt beschrieben werden. Der Aufruf DPLL(F, []) liefert entweder eine Bewertung zurück, die F erfüllt, oder UNSAT, falls F unerfüllbar ist.

```
DPLL(F, α)
  if Fα = 0
      then return UNSAT
  if Fα = 1
      then return α
  wähle x ∈ V(Fα)
  β := DPLL(F, α ∪ [x ← 0])
  if β ≠ UNSAT
      then return β
      else  return DPLL(F, α ∪ [x ← 1])
```

Die Backtracking-Strategie ist unter Umständen effizienter als ein einfaches brute-force durchprobieren aller 2^n möglichen Bewertungen: wenn für eine partielle Bewertung α bereits $C\alpha = 0$ für eine Klausel C in F ist, so ist auch $C\beta = 0$ für alle Erweiterungen $\beta \supseteq \alpha$. Diese Erweiterungen brauchen also nicht mehr untersucht werden, der Suchbaum wird an dieser Stelle gekappt und somit kleiner.

Da jeder rekursive Aufruf bei DPLL nur polynomiale Zeit benötigt, ist die Laufzeit beschränkt durch $T(n) \cdot n^{O(1)}$, wobei $T(n)$ die Größe des Rekursionsbaumes bezeichnet. Der polynomielle Faktor wird bei der Angabe der Komplexität meist unterschlagen, man gibt als Komplexität nur die Baumgröße $T(n)$ an. Ferner wird die Größe des Baumes durch die Anzahl seiner Blätter abgeschätzt, da für einen echt verzweigenden Baum die Größe durch ein konstantes Vielfaches dieser Anzahl beschränkt ist.

Man erhält so bereits eine bessere Komplexitätsschranke für k-SAT, wenn bei DPLL die folgende Strategie verfolgt wird: Wähle immer eine Klausel C, und verzweige nacheinander nach den höchstens k Variablen, die in C vorkommen.

Von den 2^k möglichen Bewertungen dieser Variablen ist immer eine α mit $C\alpha = 0$, also wird bei diesem α nicht weiter verzweigt. Der Algorithmus erzeugt also einen $(2^k - 1)$-verzweigenden Rekursionsbaum der Tiefe n/k, dieser hat $(2^k - 1)^{n/k}$ Blätter.

Die Laufzeit für k-SAT ist also beschränkt durch $O((2^k - 1)^{n/k}) = O({b_k}^n)$, wobei $b_k = \sqrt[k]{2^k - 1} < 2$ ist. Allerdings strebt b_k für große k recht schnell gegen 2, die Werte b_k für kleine k sind ungefähr:

k	3	4	5	6	7	8
b_k	1.91294	1.96799	1.98735	1.99477	1.99777	1.99903

4.1 Reduktionen

1-Klauseln

Eine besondere Rolle bei DPLL-Algorithmen spielen 1-Klauseln (engl. *unit clause*): wird nach einer Variablen x verzweigt, die in einer 1-Klausel a vorkommt, so wird der rekursive Aufruf, der $[a \leftarrow 0]$ setzt, sofort im nächsten Schritt abbrechen. Es findet also keine echte Verzweigung statt, man kann dies so auffassen, dass einfach gleich $[a \leftarrow 1]$ gesetzt wird.

Da dies für die Effizienz günstig ist, werden die 1-Klauseln in den meisten DPLL-Algorithmen bevorzugt zuerst behandelt, dies wird als *unit propagation* bezeichnet.

```
UnitProp(F, α)
  while in F gibt es eine 1-Klausel a
    F := F[a ← 1]
    α := α ∪ [a ← 1]
```

Dieser Schritt wird gesondert als Vereinfachungsschritt, eine sog. *Reduktion*, behandelt, der unmittelbar nach dem (rekursiven) Aufruf der Prozedur ausgeführt wird. Allgemein bezeichnet man als Reduktionen effizient (etwa in polynomieller Zeit) zu berechnende Abbildungen der Formeln in KNF in sich, die die Erfüllbarkeits-äquivalenz erhalten.

Neben der unit propagation gibt es weitere solche Reduktionen, darunter auch solche, die die Formel F auf andere Weise als durch Variablenbewertung modifizieren. Das allgemeine Schema eines DPLL-Algorithmus läßt sich also treffender wie folgt beschreiben:

DPLL(F, α)
  reduziere(F,α)
  if $F = 0$ then return UNSAT
  if $F = 1$ then return α
  wähle $x \in V(F)$ und $\epsilon \in \{0, 1\}$
  $\beta := \text{DPLL}(F[x \leftarrow \epsilon], \alpha \cup [x \leftarrow \epsilon])$
  if $\beta \neq$ UNSAT
    then return β
    else return DPLL($F[x \leftarrow (1 - \epsilon)], \alpha \cup [x \leftarrow (1 - \epsilon)]$)

Ein DPLL-Algorithmus hängt also von zwei Parametern ab: der Regel, nach der die Verzweigungsvariable ausgewählt wird, und den verwendeten Reduktionen. Dabei werden Reduktionen, die die Formel verkleinern, meist solange iteriert, bis keine mehr anwendbar sind.

Reine Literale

Eine fast immer verwendete Reduktion ist die Elimination von reinen Literalen (engl. *pure literals*):

Definition. *Ein Literal* a *heißt* rein *in* F*, wenn* $\bar{a}$ *nicht in* F *vorkommt.*

Ist ein Literal a rein in F, so braucht die Bewertung $[a \leftarrow 0]$ nicht betrachtet werden: Ist α eine Bewertung, die F erfüllt, mit $\alpha(a) = 0$, so wird F auch von der Bewertung α' mit $\alpha'(a) = 1$ und $\alpha'(x) = \alpha(x)$ für $x \in \text{dom}\,\alpha \setminus V(a)$ erfüllt. Man kann also reine Literale gleich mit 1 bewerten, was die folgende Reduktion rechtfertigt:

PureLit(F, α)
  while in F gibt es ein reines Literal a
    $F := F[a \leftarrow 1]$
    $\alpha := \alpha \cup [a \leftarrow 1]$

Subsumption

Eine Reduktion, die nicht durch Bewertung einer Variablen entsteht, ist die Subsumptionsregel.

Definition. *Eine Klausel* C subsumiert D*, in Zeichen* $C \leq D$*, wenn jedes Literal in* C *auch in* D *vorkommt.*

Offensichtlich erfüllt dann jede Bewertung $\alpha \models C$ auch D, also gilt die folgende Beobachtung.

Proposition 25. *Ist* F *eine Formel, in der Klauseln* C *und* D *mit* $C \leq D$ *vorkommen, so sind* F *und* $F \setminus D$ *erfüllbarkeits-äquivalent.*

Dies rechtfertigt die *Subsumptionsregel*, die aus einer Formel alle Klauseln enfernt, die von einer anderen subsumiert werden. Diese Regel gehörte zum ursprünglichen DPLL-Algorithmus, wird aber heute wegen des hohen Aufwandes meist nicht mehr verwendet.

4.2 Algorithmus von Monien-Speckenmeyer

Eine mögliche Verzweigungsstrategie ist die folgende: wähle immer die erste Variable in einer kürzesten Klausel. Im nächsten Schritt ist dann in einem der Zweige diese Klausel nicht mehr vorhanden, im Anderen ist der Rest dieser Klausel eine kürzeste Klausel. Daher kann man den DPLL-Algorithmus mit dieser Verzweigungsstrategie auch äquivalenterweise wie folgt schreiben:

simple-MS(F, α)
  if $F = 0$ then return UNSAT
  if $F = 1$ then return α
  wähle kürzeste Klausel $C = a_1 \vee \ldots \vee a_r$ in F
  for $i := 1$ to r
    $\gamma_i := [a_1 \leftarrow 0, \ldots, a_{i-1} \leftarrow 0, a_i \leftarrow 1]$
    $\beta :=$ simple-MS$(F\gamma_i, \alpha \cup \gamma_i)$
    if $\beta \neq$ UNSAT then return β
  return UNSAT

Es wird also bei Auswahl einer kürzesten Klausel r-fach verzweigt, wobei im i-ten Zweig i Variablen bewertet sind, also höchstens $n - i$ Variablen übrig bleiben.

Da stets $r \leq k$ ist, erhält man folgende Rekursionsgleichung für die Anzahl $T(n)$ der Blätter im Verzweigungsbaum

$$T(n) = T(n-1) + \ldots + T(n-k)$$

und der Ansatz $T(n) = b^n$ liefert die Gleichung $b^k = b^{k-1} + \ldots + b + 1$ für die Basis $b = b_k$. Die Lösungen dieser Gleichung für kleine Werte von k sind ungefähr:

k	3	4	5	6	7	8
b_k	1.83929	1.92757	1.96595	1.98359	1.99197	1.99603

Autarkien

Eine Verbesserung des Algorithmus läßt sich durch Verwendung des Begriffs der Autarkie erreichen.

Definition. *Eine Bewertung α heißt* autark *für* F*, wenn jede Klausel in* $F\alpha$ *auch schon in* F *vorkommt.*

Anders ausgedrückt ist α autark für F, wenn jede Klausel in F, die eine Variable in $\operatorname{dom}\alpha$ enthält, schon von α erfüllt wird. Autarke Bewertungen haben die folgende Eigenschaft:

Proposition 26. *Ist α autark für F, dann ist F erfüllbar genau dann, wenn es $\beta \models F$ gibt mit $\beta \supseteq \alpha$.*

Beweis. Ist $F\alpha$ erfüllbar, so ist offensichtlich F erfüllbar. Umgekehrt sei $\beta \models F$, dann definiere $\beta' \supseteq \alpha$ durch

$$\beta'(x) = \begin{cases} \alpha(x) & x \in \operatorname{dom}\alpha \\ \beta(x) & x \in \operatorname{dom}\beta \setminus \operatorname{dom}\alpha \end{cases}$$

Jede Klausel in F, die durch α erfüllt wird, wird auch durch β' erfüllt. Die übrigen Klauseln in F, also diejenigen in $F\alpha$, enthalten die Variablen in $\operatorname{dom}\alpha$ nicht, müssen also von β dadurch erfüllt werden, dass Literale aus Variablen in $\operatorname{dom}\beta \setminus \operatorname{dom}\alpha$ mit 1 bewertet werden. Diese Klauseln werden also ebenfallso durch β' erüllt, also gilt $\beta' \models F$. □

Die Verbesserung des Algorithmus besteht darin, dass nach Auswahl einer kürzesten Klausel erst überprüft wird, ob eine der möglichen erfüllenden Bewertungen dieser Klausel autark ist. Ist dies der Fall, wird nur der Zweig für diese Bewertung weiterverfolgt, andernfalls wird wiederum über alle möglichen Bewertungen verzweigt.

```
MS(F, α)
  if F = 0 then return UNSAT
  if F = 1 then return α
  wähle kürzeste Klausel C = a_1 ∨ ... ∨ a_r in F
  for i := 1 to r
      γ_i := [a_1 ← 0, ..., a_{i-1} ← 0, a_i ← 1]
      if γ_i autark für F then return MS(Fγ_i, α ∪ γ_i)
  for i := 1 to r
      β := MS(Fγ_i, α ∪ γ_i)
      if β ≠ UNSAT then return β
  return UNSAT
```

Ist nun keine der Bewertungen $\gamma_1, \ldots, \gamma_r$ autark, so muss es für jede dieser Bewertungen γ_i eine Klausel C' geben mit $C'\gamma_i \neq 1$, aber $V(C') \cap \operatorname{dom} \gamma_i \neq \emptyset$. Es wird also ein Literal in C' mit 0 bewertet, also hat die Klausel $C'\gamma_i$ höchstens $k-1$ Literale. Für jedes $i = 1, \ldots, r$ enthält also $F\gamma_i$ eine Klausel mit höchstens $k-1$ Literalen, daher wird im nächsten Schritt höchstens $(k-1)$-fach verzweigt.

Bezeichnen wir wiederum mit $T(n)$ die Größe des Verzweigungsbaumes für Formeln in n Variablen, und mit $T'(n)$ die Größe des Verzweigungsbaumes für Formeln in n Variablen, die mindestens eine Klausel der Breite höchstens $k-1$ enthalten. Dann gelten die folgenden Rekursionsgleichungen:

$$\begin{aligned} T(n) &= \max(T(n-1), T'(n-1) + \ldots + T'(n-k)) \\ T'(n) &= \max(T(n-1), T'(n-1) + \ldots + T'(n-k+1)) \end{aligned}$$

Wir zeigen nun, dass für alle n gilt

$$T(n) \leq T'(n) + T'(n-1) \,,$$

also wird in den obigen Gleichungen das Maximum stets im zweiten Term angenommen. Wir setzen für $T(n)$ und $T'(n)$ die Definition ein, und definieren $t_1 := T'(n-1) + \ldots + T'(n-k)$ und $t_2 := T'(n-1) + \ldots + T'(n-k+1)$. Dann ist zu zeigen

$$\max(T(n-1), t_1) \leq \max(T(n-1), t_2) + T'(n-1) \qquad (*)$$

Da $t_1 = t_2 + T'(n-k)$ ist, gilt offenbar $t_2 \leq t_1$.

Falls nun $T(n-1) \geq t_1$ ist, gilt auch $T(n-1) \geq t_2$, also wird auf beiden Seiten das Maximum im ersten Term angenommen, daher gilt $(*)$, da $T(n-1) \leq T(n-1) + T'(n-1)$ ist.

Andernfalls ist $T(n) = \max(T(n-1), t_1) = t_1$, und es gilt

$$\begin{aligned} t_1 &= t_2 + T'(n-k) \\ &\leq \max(T(n-1), t_2) + T'(n-k) \\ &\leq T'(n) + T'(n-2) + \ldots + T'(n-k) \\ &\leq T'(n) + \max(T(n-2), T'(n-2) + \ldots + T'(n-k)) \\ &= T'(n) + T'(n-1) \end{aligned}$$

also gilt $(*)$ auch in diesem Fall.

Die Rekursionsgleichungen vereinfachen sich somit zu

$$\begin{aligned} T(n) &= T'(n-1) + \ldots + T'(n-k) \\ T'(n) &= T'(n-1) + \ldots + T'(n-k+1) \end{aligned}$$

und der Ansatz $T'(n) = b^n$ ergibt die Gleichung $b^{k-1} = b^{k-2} + \ldots + b + 1$ für die Basis $b = b_k$. Die erste Gleichung liefert auch $T(n) \leq b_k^{n-1} + \ldots + b_k^{n-k} = O({b_k}^n)$, somit ist die Komplexität des Algorithmus im wesentlichen $O({b_k}^n)$, wobei die Werte b_k für kleine Werte von k die folgenden sind:

k	3	4	5	6	7	8
b_k	1.61804	1.83929	1.92757	1.96595	1.98359	1.99197

Eine Verfeinerung dieser Methode für 3-SAT wurde von Schiermeyer angegeben, die eine Komplexität von b^n mit $b = 1.5782$ erreicht.

4.3 Analysemethode von Kullmann

Für einen Vektor $(d_1, \ldots, d_m)$ mit $m \geq 1$ und $d_i > 0$ für $i = 1, \ldots, m$ sei $\tau(d_1, \ldots, d_m)$ definiert als die eindeutig bestimmte positive Lösung der Gleichung

$$x^{-d_1} + \ldots + x^{-d_m} = 1 \, .$$

Sei T ein Baum, und d eine Kantenbewertung: jedem Knoten v mit Sohn w wird eine Distanz $d(v, w) > 0$ zugeordnet. Für einen Knoten v mit den Söhnen $w_1, \ldots, w_m$ sei $d(v) = (d_1, \ldots, d_m)$ mit $d_i := d(v, w_i)$für $i = 1, \ldots, m$. Dann ist die *Verzweigungszahl* $\tau(v)$ von v definiert als $\tau(d(v))$, und $\tau(T) := \max(\tau(v))$ über alle inneren Knoten v in T.

Für einen Pfad $p = v_1, \ldots, v_m$ in T sei ferner

$$d(p) = \sum_{i=1}^{m-1} d(v_i, v_{i+1})$$

und $d(T) := \max(d(p))$ über alle Pfade in T von der Wurzel zu einem Blatt. Dann gilt die folgende obere Schranke an die Größe des Baumes:

Satz 27. *Die Anzahl der Blätter in* T *ist höchstens* $\tau(T)^{d(T)}$.

Beweis. Für jeden inneren Knoten gilt $\tau(v)^{d_1} + \ldots + \tau(v)^{d_m} = 1$, außerdem ist $\tau(v) \geq 1$, und und $\tau(v) = 1$ genau dann, wenn $m = 1$ ist, und somit ist $0 < \tau(v)^{-d_i} \leq 1$ für alle i. Daher bilden die Werte $p_i := \tau(v)^{-d_i}$ eine Wahrscheinlichkeitsverteilung auf den Kanten, die wir so interpretieren, dass von v mit Wahrscheinlichkeit p_i zum Sohn w_i übergegangen wird.

Dadurch wird ein zufälliger Pfad in T von der Wurzel zu einem Blatt erzeugt. Sei w ein Blatt, und $p = v_1, \ldots, v_m = w$ der eindeutig bestimmte Pfad von der Wurzel zu w. Die Wahrscheinlichkeit, dass der zufällige Pfad

w erreicht, ist genau die Wahrscheinlichkeit, dass er gerade p ist. Diese Wahrscheinlichkeit ist beschränkt durch

$$\prod_{i=1}^{m-1} \tau(v_i)^{-d(v_i,v_{i+1})} \geq \prod_{i=1}^{m-1} \tau(T)^{-d(v_i,v_{i+1})}$$
$$\geq \tau(T)^{-\sum_{i=1}^{m-1} d(v_i,v_{i+1})} = \tau(T)^{-d(p)} \geq \tau(T)^{-d(T)}$$

also wird jedes Blatt mit Wahrscheinlichkeit $\tau(T)^{-d(T)}$ erreicht. Daher kann die Anzahl der Blätter höchstens $\tau(T)^{d(T)}$ sein. □

Um eine möglichst gute Analyse der Laufzeit eines DPLL-Verfahrens in Abhängigkeit von n zu erhalten, gilt es also, eine Distanzfunktion d für den Rekursionsbaum T zu finden, derart dass $d(T) \leq n$ ist, und $\tau(T)$ möglichst klein wird.

Um zu sehen, wie diese Methode die bisher verwendete verallgemeinert, präsentieren wir hier noch einmal die Analyse von simple-MS für 3-SAT in diesem Rahmen:

Definiere die Distanz $d(v, w)$ als die Anzahl der Variablen, die beim Übergang von v nach w bewertet werden. Auf jedem Pfad können höchstens alle Variablen bewertet werden, also ist für jeden Pfad $d(p) \leq n$, also $d(T) \leq n$.

Jeder innere Knoten im Rekursionsbaum T hat drei Söhne, je einen mit den Distanzen 3, 2 und 1. Also ist $\tau(v) = \tau(3, 2, 1)$ die Lösung der Gleichung

$$x^{-3} + x^{-2} + x^{-1} = 1$$

und da wir nur an positiven Lösungen interessiert sind, können wir diese mit x^3 multiplizieren. Wir erhalten dieselbe Gleichung

$$x^3 = x^2 + x + 1$$

die wir bei der obigen Analyse von gelöst haben, also ist $\tau(T)$ ungefähr 1.83929, und wir erhalten das gleiche Ergebnis wie oben.

Zur Analyse des Algorithmus von Monien-Speckenmeyer für 3-SAT wird dieselbe Distanzfunktion wie oben benutzt. Hier gibt es drei verschiedene Typen von inneren Knoten:

- Hat ein Knoten v genau einen Sohn (im Fall, dass eine autarke Bewertung gefunden wurde), so ist $\tau(v) = \tau(1) = 1$.
- Für Knoten v mit zwei Söhnen hat einer die Distanz 2, der andere Distanz 1, also ist $\tau(v) = \tau(2, 1)$ die positive Lösung von

  $$x^{-2} + x^{-1} = 1 \quad \text{bzw.} \quad x^2 = x + 1\,,$$

  also der goldene Schnitt $\phi \approx 1.61804$.

- Daneben gibt es Knoten v mit drei Söhnen, je einen mit den Distanzen 3, 2 und 1. Für diese ist wie oben $\tau(v) = \tau(3,2,1)$.

Es scheint also gegenüber dem vereinfachten Algorithmus zunächst nichts gewonnen zu sein.

Es gilt jedoch, dass ein Knoten v — mit Ausnahme der Wurzel — nur dann drei Söhne haben kann, wenn er keine 2-Klausel enthält, also durch eine autarke Bewertung erzeugt wurde. Dann ist v der einzige Sohn seines Vaters v', und hat von diesem mindestens die Distanz 1.

Um dies bei der Analyse auszunutzen, können v und v' zu einem neuen Knoten v^* verschmolzen werden. Dieser hat nun drei Söhne, jeweils mit den Distanzen 4, 3 und 2, also ist $\tau(v^*) = \tau(4,3,2)$ bestimmt durch Gleichung

$$x^{-4} + x^{-3} + x^{-2} = 1 \quad \text{bzw.} \quad x^4 = x^2 + x + 1$$

deren positive Lösung ist ungefähr $\tau(4,3,2) < 1.46558 < \tau(2,1)$. Damit ist $\tau(T) = \tau(2,1) < 1.61804$, und wir erhalten also obere Schranke für die Größe des Baumes $O(1.61804^n)$.

Eine alternative Analyse, bei der die Struktur des Baumes erhalten bleibt, besteht darin, die Distanz zwischen v und v' um einen Wert $\epsilon < 1$ zu vermindern, und die Distanzen zwischen v und seinen drei Söhnen um den gleichen Betrag zu erhöhen. Dadurch ändern sich die Pfadlängen $d(p)$ im Baum nicht, und für v erhalten wir die neue Verzweigungszahl $\tau(v) = \tau(3+\epsilon, 2+\epsilon, 1+\epsilon)$. Für $\epsilon = 1/2$ ist dies $\tau(7/2, 5/2, 3/2) < 1.59704 < \tau(2,1)$, also erhalten wir wiederum $\tau(T) = \tau(2,1)$.

Offensichtlich spielt die Reihenfolge der Argumente für den Wert der τ-Funktion keine Rolle. Weiterhin hat sie die folgenden Eigenschaften, von denen wir einige oben schon ausgenutzt haben:

Proposition 28. *1. Ist* $d_1' > d_1$, *so ist* $\tau(d_1', \ldots, d_m) < \tau(d_1, \ldots d_m)$.

2. Ist $d_1' + d_2' = d_1 + d_2$, *und* $\min(d_1', d_2') \geq \min(d_1, d_2)$, *dann ist*

$$\tau(d_1', d_2', d_3, \ldots, d_m) \leq \tau(d_1, d_2, d_3, \ldots, d_m)$$

wobei Gleichheit nur dann gilt, wenn sie in der Voraussetzung gilt.

3. Sei $d = (d_1, \ldots, d_m)$ *und* $d' := (d_1', \ldots, d_n')$, *und sei* $d^* := (d_1 + d_1', \ldots, d_1 + d_n', d_2, \ldots, d_m)$. *Dann gilt:*

- *Ist* $\tau(d) \leq \tau(d')$, *so ist* $\tau(d) \leq \tau(d^*) \leq \tau(d')$.
- *Ist* $\tau(d) \geq \tau(d')$, *so ist* $\tau(d) \geq \tau(d^*) \geq \tau(d')$.

Dabei sind beide Ungleichungen echt, falls die Ungleichung in der Voraussetzung echt ist.

Anschaulich haben diese Eigenschaften die folgende Bedeutung, die man bei Entwurf und Analyse von DPLL-Verfahren ausnutzen kann:

1. Vergrößern der Distanz zu einem Sohn verbessert die Verzweigungszahl eines Knotens. Beispielsweise ist $\tau(3,1) < 1.46558 < \tau(2,1)$.
2. Von zwei Verzweigungen mit gleicher Summe der Distanzen liefert die symmetrischere die bessere Verzweigungszahl. Beispielsweise ist $\tau(2,2) = \sqrt{2} < \tau(3,1)$.
3. Werden zwei Knoten zu einem verschmolzen, so liegt die Verzweigungszahl des neuen Knotens zwischen denen der alten Knoten. Beispielsweise ist $\tau(4,3,2) = \tau(3,1+3,1+1) = \tau(3,1)$, und

$$\tau(3,1) < \tau(3,3,2) = \tau(3,1+2,1+1) < 1.52138 < \tau(2,1)\ .$$

4.4 Der Algorithmus von Zhang

Bei der Analyse des Algorithmus von Monien-Speckenmeyer für 3-SAT hat sich gezeigt, dass es für die Komplexität vorteilhaft ist, wenn in der Formel 1-Klauseln oder 2-Klauseln vorkommen.
Zur Beschreibung und Analyse des Algorithmus wird die Formel F stets partitioniert in $F = U \cup D \cup T \cup F'$ wie folgt:

- U ist die Menge der 1-Klauseln in F
- D ist eine maximale Menge von 2-Klauseln, die mit den Klauseln in U und untereinander keine gemeinsamen Variablen haben.
- T ist die Menge der übrigen 2-Klauseln.

Wir definieren daraus die Parameter $u := |U|$, $d := |D|$ und $t := \min(|T|, 2)$.
Der Algorithmus von Zhang nutzt den Vorteil durch die Existenz kurzer Klauseln aus, indem er sicherstellt, dass stets kurze Klauseln, also 1- oder 2-Klauseln vorhanden sind. Dabei nimmt man o.E. an, dass die Eingabeklausel F bereits mindestens eine kurze Klausel enthält.
Um dies zu ermöglichen, wird auf Reduktionen, die unkontrolliert kurze Klauseln eliminieren (wie z.B. das Beseitigen reiner Literale), verzichtet. Die einzigen verwendeten Reduktionen sind:

- Teste, ob komplementäre 1-Klauseln a und $\bar{a}$ vorhanden sind, in diesem Fall gib UNSAT zurück.

- Lösche subsumierte 3-Klauseln.
- Falls $u \geq 2$ oder $u = 1$ und $d > 0$ ist, wähle eine 1-Klausel a und setze $[a \leftarrow 1]$.

Die andere wesentliche Idee ist, neben autarken Bewertungen, durch die keine neuen Klauseln entstehen, auch solche Bewertungen gesondert zu behandeln, bei denen nur eine Variable bewertet wird, und durch die nur eine neue Klausel entsteht.

Definition. *Eine Bewertung* α *mit* $|\operatorname{dom}\alpha| = 1$ *heißt* quasi-autark *für* F, *wenn* $|F\alpha \setminus F| = 1$ *ist.*

Das heißt in anderen Worten, die Bewertung $[a \leftarrow 1]$ ist quasi-autark, wenn a ein fast reines Literal ist, also $\bar{a}$ nur einmal in F vorkommt.

Der Algorithmus von Zhang wird beschrieben durch die folgende rekursive Prozedur.

```
Zh(F, α)
  (F, α) := reduziere(F, α)
  if F = 0 then return UNSAT
  if F = 1 then return α
  if u = 1 then return unit(F, α)
  (γ1, γ2) := verzweige(F, α)
  if γi autark für F für ein i = 1, 2
       then return aut(F, γi, α)
  if γi quasi-autark für F für ein i = 1, 2
       then return qu-aut(F, γi, α)
  β := Zh(Fγ1, α ∪ γ1)
  if β ≠ UNSAT
       then return β
       else return Zh(Fγ2, α ∪ γ2)
```

Nach Reduktion ist entweder $u = 0$, oder es ist $u = 1$ und $d = 0$, d.h. es ist noch eine 1-Klausel a vorhanden, und jede 2-Klausel enthält entweder a oder $\bar{a}$. In diesem Fall wird wie folgt vorgegangen:

```
unit(F, α)
  wähle eine 3-Klausel (b ∨ c ∨ d)
  γ := [a ← 1, b ← 0, c ← 0, d ← 1]
  if γ autark für F
       then return aut(F, γ, α)
  β := Zh(F[a ← 1] ∧ (b ∨ c), α ∪ [a ← 1])
```

if $\beta \neq \text{UNSAT}$
then return β
else return $\text{Zh}(F\gamma, \alpha \cup \gamma)$

Sind keine 1-Klauseln mehr vorhanden ($u = 0$), werden Bewertungen γ_1, γ_2 gemäss der folgenden Fallunterscheidung berechnet und nach diesen verzweigt.

Fall 1: $t \geq 1$, d.h. es gibt 2-Klauseln mit gemeinsamen Variablen.

Fall 1.1: es gibt 2-Klauseln $(a \vee b)$ und $(\bar{a} \vee c)$.

$\gamma_1 := [a \leftarrow 1, c \leftarrow 1]$
$\gamma_2 := [a \leftarrow 0, b \leftarrow 1]$

Fall 1.2: andernfalls wähle 2-Klauseln $(a \vee b)$ und $(a \vee c)$.

$\gamma_1 := [a \leftarrow 1]$
$\gamma_2 := [a \leftarrow 0, b \leftarrow 1, c \leftarrow 1]$

Fall 2: $t = 0$, d.h. alle 2-Klauseln sind variablen-disjunkt.

Wähle eine 2-Klausel $(a \vee b)$, und setze

$\gamma_1 := [a \leftarrow 1]$
$\gamma_2 := [a \leftarrow 0, b \leftarrow 1]$

Beim Auftreten einer autarken Bewertung γ wird nicht verzweigt, sondern nur γ weiter verfolgt, dabei wird aber dafür gesorgt, dass eine 1-Klausel in der Formel entsteht.

$\text{aut}(F, \gamma, \alpha)$
sei $\gamma = \gamma' \cup [a \leftarrow 1]$
return $\text{Zh}(F\gamma' \wedge a, \alpha \cup \gamma')$

Zur Behandlung quasi-autarker Bewertungen benötigen wir die folgenden zwei Lemmas. Ist F eine Formel, a ein Literal in F und C eine Klausel, dann bezeichnet $F\langle a := C\rangle$ die Formel, die aus F ensteht, indem jedes Vorkommen von a durch die Klausel C ersetzt wird. Die Vorkommen von $\bar{a}$ bleiben dabei unangetastet.

Lemma 29. *Falls* a *in* F *nicht vorkommt, dann sind* $F \wedge (a \vee C)$ *und* $F\langle \bar{a} := C\rangle$ *erfüllbarkeits-äquivalent.*

Beweis. Sei $\alpha \models F \wedge (a \vee C)$. Ist $\alpha(a) = 0$, so muss $\alpha(\bar{a}) = \alpha(C) = 1$ sein, also ist $F\langle \bar{a} := C\rangle\alpha = F\alpha = 1$.
Ist andererseits $\alpha(a) = 1$, so ist ist in jeder Klausel D in F, die $\bar{a}$ enthält, ein anderes Literal erfüllt, also ist $D\langle \bar{a} := C\rangle\alpha = D\alpha = 1$, und somit $F\langle \bar{a} := C\rangle\alpha = 1$. Also gilt in jedem Fall $\alpha \models F\langle \bar{a} := C\rangle$.

Sei umgekehrt $\alpha \models F\langle \bar{a} := C\rangle$, und o.E. a von α unbewertet. Ist $\alpha \models C$, dann setze $\alpha' := \alpha \cup [a \leftarrow 0]$, dann folgt $\alpha \models F \wedge (a \vee C)$.

Ist andererseits $C\alpha = 0$, dann muss in jeder Klausel D in $F\langle \bar{a} := C\rangle$ ein Literal, dass nicht in C ist, erfüllt sein, also gilt auch $\alpha \models F$. Also gilt $\alpha' \models F \wedge (a \vee C)$, wobei $\alpha' := \alpha \cup [a \leftarrow 1]$. □

Korollar 30. *Falls* a *in* F *nur in der Klausel* $(a \vee b \vee C)$ *vorkommt, dann ist* F *genau dann erfüllbar, wenn eine der folgenden Formeln erfüllbar ist.*

$$F[b \leftarrow 1, a \leftarrow 0]$$
$$F\langle \bar{a} := C\rangle[b \leftarrow 0, a \leftarrow 1]$$

Beweis. In $F[b \leftarrow 1]$ kommt a nicht mehr vor, also ist es erfüllbarkeits-äquivalent zu $F[b \leftarrow 1, a \leftarrow 0]$.

In $F[b \leftarrow 0]$ kommt a nur in der Klausel $a \vee c$ vor also ist es nach Lemma 29 erfüllbarkeits-äquivalent zu $F\langle \bar{a} := c\rangle[b \leftarrow 0]$. Da aber nun $\bar{a}$ nicht mehr vorkommt, ist diese erfüllbarkeits-äquivalent zu $F\langle \bar{a} := c\rangle[b \leftarrow 0, a \leftarrow 1]$. □

Lemma 31. *Falls* a *und* b *in* F *nicht vorkommen, dann ist* $F \wedge (a \vee b \vee c)$ *genau dann erfüllbar, wenn eine der folgenden Formeln erfüllbar ist.*

$$F[c \leftarrow 1, a \leftarrow 0, b \leftarrow 0]$$
$$F[c \leftarrow 0, a \leftarrow 1, b \leftarrow 0]$$
$$F[c \leftarrow 0, a \leftarrow 0, b \leftarrow 1]$$

Dabei können die komplementären Literale $\bar{a}$ und $\bar{b}$ durchaus in F vorkommen.

Beweis. Da a und b in F nicht vorkommen, kann jedes $\alpha \models F$ so abgeändert werden, dass a oder b mit 0 bewertet werden.

Ist $\alpha(c) = 1$, so gilt $\alpha \models (a \vee b \vee c)$, und daher können a und b mit 0 bewertet werden. Ist andererseits $\alpha(c) = 1$, so muss eines von a und b mit 1 bewertet werden, es genügt aber auch eines. □

Beim Auftreten einer quasi-autarken Bewertung wird die Verzweigung verworfen und wie folgt eine neue berechnet. Ist $[a \leftarrow 1]$ quasi-autark, so kommt $\bar{a}$ in F nur in einer Klausel C vor. Diese kann nur eine 3-Klausel $C = (\bar{a} \vee b \vee c)$ sein, da Quasi-autarkie nur in den Fällen 1.2 oder 2 auftreten kann. Wäre eine 2-Klausel $(\bar{a} \vee b)$ in F vorhanden, so wäre aber Fall 1.1 eingetreten.

Nach Korollar 30 ist also F genau dann erfüllbar, wenn eine der Formeln $F[b \leftarrow 1, a \leftarrow 1]$ oder $F\langle a := c\rangle[b \leftarrow 0, a \leftarrow 0]$ erfüllbar ist. Es wird dann unterschieden, on eine dieser Formeln keine neuen Klauseln enthält.

Enthält $F\langle a := c\rangle[b \leftarrow 0, a \leftarrow 0]$ keine neuen Klauseln, so kommt b in F nur in der Klausel C vor. Also ist nach Lemma 31 die in der folgenden Prozedur gewählte Verzweigung korrekt.

```
qu-aut(F, γ, α)
  sei γ = [a ← 1]
  sei Fγ \ F = {(b ∨ c)}
  if [b ← 1, a ← 1] ist autark für F
      then return aut(F, [b ← 1, a ← 1], α)
  if F⟨a := c⟩[b ← 0, a ← 0] ⊆ F
      then γ1 := [c ← 1, a ← 1, b ← 0]
           γ2 := [c ← 0, a ← 0, b ← 0]                    (A)
           γ3 := [c ← 0, a ← 1, b ← 1]
           if es gibt 2-Klausel (b̄ ∨ d) mit d ∉ {a, ā, c, c̄}
               then γ3 := γ3 ∪ [d ← 1]                    (B)
           if γi autark für F für ein i = 1, ..., 3
               then return aut(F, γi, α)
           for i := 1 to 3
               β := Zh(Fγi, α ∪ γi)
               if β ≠ UNSAT then return β
           return UNSAT
  β := Zh(F[b ← 1, a ← 1], α ∪ [b ← 1, a ← 1])            (C)
  if β ≠ UNSAT
      then return β
      else return Zh(F⟨a := c⟩[b ← 0, a ← 0], α ∪ [b ← 0, a ← 0])
```

Verwendet man zur Analyse des Algorithmus als Distanzfunktion wie oben die Anzahl der bewerteten Variablen, so erhält man keine bessere Schranke als die für den Algorithmus von Monien-Speckenmeyer, da im Fall 2 bei der Verzweigung Knoten auftreten, die zwei Söhne mit Distanzen 2 und 1 haben. Wir verwenden daher eine feinere Distanzfunktion, die die Anzahl der kurzen Klauseln mit berücksichtigt.

Sei ϵ ein Parameter, dessen Wert wir später festlegen werden. Als Maß für die Komplexität einer Formel F definieren wir

$$\mu := n - \epsilon(u + d + t) \; .$$

Ist v ein Knoten im Rekursionsbaum, und F die bei v betrachtete Formel mit Maß μ, sowie w ein Sohn von v und F' die bei w betrachtete Formel mit Maß μ', so definieren wir

$$d(v, w) = \mu - \mu'$$

Die Distanz zweier Knoten ist also die Anzahl der bewerteten Variablen plus der (gewichteten) Differenz der Werte u, d und t, die Auskunft über die kurzen Klauseln in der Formel geben.
An den folgenden Stellen im Algorithmus treten Verzweigungen auf:

- Prozedur unit()

  Im ersten Zweig verschwindet eine 1-Klausel, alle evtl. vorhandenen 2-Klauseln werden zu 1-Klauseln, und es wird eine 2-Klausel hinzugefügt. Daher bleibt $(u + d + t)$ konstant, die Distanz ist 1.

  Im zweiten Zweig verschwindet eine 1-Klausel, schlimmstenfalls alle 2-Klauseln, aber es wird eine neue 2-Klausel erzeugt, da γ nicht autark ist wenn echt verzweigt wird. Die Distanz ist also mindestens $4 - 2\epsilon$.

- Fall 1.1

  Im schlimmsten Fall verschwinden in beiden Zweigen zwei Klauseln aus D und alle aus T, da aber (im Fall einer echten Verzweigung) keine Autarkie vorliegt, ensteht mindestens eine neue kurze Klausel. Daher nimmt $(u + d + t)$ höchstens um 3 ab, die Distanz ist also mindestens $2 - 3\epsilon$.

- Fall 1.2

  Im ersten Zweig verschwindet nur eine Klausel aus D, aber möglicherweise alle aus T. Da aber γ_1 nicht quasi-autark ist, kommen auch zwei neue kurze Klauseln hinzu. Die Distanz ist also mindestens $1 - \epsilon$.

  Im zweiten Zweig verschwinden höchstens zwei Klauseln aus D und möglicherweise alle aus T, aber es kommt eine kurze Klausel hinzu, da γ_2 nicht autark ist. Somit ist die Distanz mindetens $3 - 3\epsilon$.

- Fall 2

  In jedem der Zweige verschwindet nur eine Klausel aus D. Im ersten Zweig kommen zwei neue kurze Klauseln, da γ_1 nicht quasi-autark ist, also ist die Distanz mindestens $1 + \epsilon$ Im zweiten kommt mindestens eine kurze Klausel hinzu, also ist die Distanz mindestens 2.

- Prozedur qu-aut(), Fall (A)

  Falls $\bar{b}$ nicht, oder nur gemeinsam mit $a, \bar{a}, c, \bar{c}$ in 2-Klauseln vorkommt, so verschwinden höchstens zwei Klauseln aus D und schlimmstenfalls alle aus T, es kommt aber eine neue 2-Klausel hinzu. Daher ist die Distanz in jedem Zweig mindestens $3 - 3\epsilon$.

- Prozedur qu-aut(), Fall (B)

  Kommt dagegen $\bar{b}$ gemeinsam mit einem anderen Literal d in einer 2-Klausel vor, so können drei Klauseln aus D verschwinden. Daher wird in einem Zweig eine vierte Variable eliminiert, so dass die Distanz in zwei Zweigen mindestens $3-4\epsilon$ und im dritten mindestens $4-4\epsilon$ beträgt.

- Prozedur qu-aut(), Fall (C)

  Dieser Fall ist analog zum Fall 1.1.

Der Wert $\tau(T)$ ist dementsprechend das Maximum aus den folgenden Werten, die Knoten entsprechen, wo gemäß der genannten Fälle verzweigt wird.

$\tau(1, 4-2\epsilon)$	Prozedur unit()
$\tau(2-3\epsilon, 2-3\epsilon)$	Fall 1.1 und Prozedur qu-aut(), Fall (C)
$\tau(1-\epsilon, 3-3\epsilon)$	Fall 1.2
$\tau(1+\epsilon, 2)$	Fall 2
$\tau(3-3\epsilon, 3-3\epsilon, 3-3\epsilon)$	Prozedur qu-aut(), Fall (A)
$\tau(3-4\epsilon, 3-4\epsilon, 4-4\epsilon)$	Prozedur qu-aut(), Fall (B)

Der Wert ϵ, bei dem das Maximum dieser Werte minimal wird, liegt ungefähr bei $\epsilon = 0,1528477$, und für diesen ist $\tau(T) = \tau(1+\epsilon, 2) < 1.570214$. Daher ist die Größe des Rekursionsbaumes, und somit die Komplexität des Algorithmus ungefähr $O(1.570214^n)$.

Satz 32. *Zhangs Algorithmus löst 3-SAT in Zeit* $O(1.570214^n)$.

Ähnliche, noch komplexere Algorithmen für 3-SAT wurden in der Folge angegeben, von Kullmann mit einer Laufzeit $O(1.5045^n)$ und Schiermeyer mit einer Laufzeit von $O(1.4963^n)$.

6 Probabilistische Algorithmen

In diesem Kapitel werden probabilistische Algorithmen betrachtet, also solche, die zufällige Entscheidungen treffen. Typischerweise sind solche Algorithmen unvollständig, d.h., sie finden für erfüllbare Formeln mit hoher Wahrscheinlichkeit eine erfullende Bewertung, können aber nicht definitiv die Unerfüllbarkeit einer Formel feststellen.

Eine Klasse solcher Algorithmen verwendet eine lokale Verbesserungsstrategie. Diese Algorithmen arbeiten nach dem folgenden Schema:

```
wähle eine totale Bewertung α
repeat
    if α ⊨ F then return α
    wähle Variable x
    α := α with [x ← 1 − α(x)]
```

Dies wird solange iteriert, bis entweder eine erfüllende Bewertung $\alpha \models F$ gefunden wird, oder bis hinreichend lange gesucht wurde, dass man mit einiger Sicherheit auf die Unerfüllbarkeit von F geschlossen werden kann.

Verschiedene Algorithmen unterscheiden sich darin, wie einerseits die Anfangsbewertungen α gewählt werden, und wie die Variable x, an der die Bewertung bei der Suche geändert wird, ausgewählt wird. Beide Auswahlen können deterministisch oder zufällig getroffen werden.

Werden die Stellen x, an der die Bewertung geändert wird, systematisch ausgewählt, aber die Anfangsbewertungen zufällig, so führt dies zu einem einfachen probabilistischen Algorithmus, dem *Hammingkugel*-Algorithmus. Eine derandomisierte Version davon, bei der auch die Anfangsbewertungen deterministisch ausgewählt werden, führt zu einem deterministischen Algorithmus für k-SAT: er erreicht bei 3-SAT die Komplexität $O(1.5^n)$, und kann durch geschickte Organisation der Suche so verbessert werden, dass er eine Laufzeit von $O(1.473^n)$ erreicht, und damit eine bessere Laufzeit erreicht als die DPLL-basierten Algorithmen aus Kapitel 3.

Werden dagegen beide Auswahlen zufällig getroffen, führt dies zum Irrfahrt-Algorithmus (*random walk*). Dieser war mit einer Komplexität von $O((4/3)^n)$ der zum Zeitpunkt seiner Veröffentlichung beste probabilistische Algorithmus für 3-SAT. Eine noch leicht verbesserte Version erreicht die Laufzeit $O(1.3302^n)$.

Schließslich betrachten wir noch einen probabilistischen Algorithmus von Paturi, Pudlák und Zane, der der Vorgehensweise bei DPLL ähnlicher ist als diese Algorithmen. Eine verbesserte Version dieses Algorithmus von Paturi, Pudlák, Saks und Zane ist derzeit noch immer der asymptotisch beste

probablilistische Algorithmus für k-SAT.

6.1 Hammingkugel-Algorithmus

Für zwei Bewertungen α, β ist die Hamming-Distanz

$$d(\alpha, \beta) := \left|\{x; \alpha(x) \neq \beta(x)\}\right| ,$$

die Hamming-Kugel vom Radius r um α ist

$$H(\alpha, r) := \{\, \beta \,;\, d(\alpha, \beta) \leq r \,\} .$$

Offenbar ist $|H(\alpha, r)| =: V(r) = \sum_{i=0}^{r} \binom{n}{r}$ unabhängig von α. Sei $\rho := r/n$ der normierte Radius. Es ist eine aus der Informationstheorie bekannte Tatsache, dass für $0 < \rho \leq 1/2$ gilt

$$\frac{1}{\sqrt{8n\rho(1-\rho)}} 2^{h(\rho)n} \leq V(r) \leq 2^{h(\rho)n} ,$$

wobei $h(\rho) := -\rho \log \rho - (1-\rho)\log(1-\rho)$ die Entropiefunktion ist. Ist also ρ konstant, so ist bis auf einen polynomiellen Faktor $V(r) \approx 2^{h(\rho)n}$.

Der folgende Algorithmus durchsucht die Hammingkugel $H(\alpha, r)$ nach einer erfüllenden Bewertung.

```
suche(α, r)
  if α ⊨ F
     then return α
  if r = 0
     then return NIL
  wähle C = a_1 ∨ ... ∨ a_k mit Cα = 0
  for i := 1 to k do
     α_i := α with [a_i ← 1]
     β := suche(α_i, r − 1)
     if β ≠ NIL
        then return β
  return NIL
```

Der Zeitaufwand ist offenbar k^r, da ein k-fach verzweigender Suchbaum der Tiefe r aufgespannt wird. Dies ist unter Umständen weniger als das Volumen $V(r)$. So ist z.B. für $k = 3$ und $r = n/2$ das Volumen $V(r) \geq 2^{n-1}$, aber $k^r = \sqrt{3}^n \geq 1.7321^n$.

Damit ergibt sich ein einfacher deterministischer Algorithmus für 3-SAT, der für die Wertzuweisungen α_0, die alle Variablen mit 0 bewertet, und

α_1, die alle mit 1 bewertet, nacheinander suche$(\alpha_0, n/2)$ und suche$(\alpha_0, n/2)$ aufruft. Die Komplexität ist $O(\sqrt{3}^n) = O(1.7321^n)$.

Eine bessere Komplexität erreicht man, wenn man einen kleineren Radius $r = \rho n$ für ein $\rho < \frac{1}{2}$ wählt, und dafür mehrere Startpunkte zufällig wählt. Damit erhält man die folgende probabilistische Variante des Hammingkugel-Algorithmus:

```
for i := 1 to w do
      wähle α_i zufällig
      β := suche(α_i, ρn)
      if β ≠ NIL
          then return β
  return NIL
```

Dabei werden die Anfangsbewertungen α_i gleichverteilt und unabhängig aus den 2^n möglichen Bewertungen der n Variablen gezogen.

Die Komplexität ist offenbar $w \cdot k^{\rho n}$. Es gilt nun, die Parameter w und ρ so zu wählen, dass die Fehlerwahrscheinlichkeit vernachlässigbar klein wird, und dabei die Laufzeit minimiert wird.

Der einzige mögliche Fehler ist, dass die Eingabeformel erfüllbar ist, aber der Algorithmus keine erfüllende Bewertung findet.

Sei also $\alpha \models F$. Ein Aufruf von suche$(\alpha_i, \rho n)$ findet α, falls $\alpha_i \in H(\alpha, r)$ ist, was mit Wahrscheinlichkeit $V(\rho n)2^{-n}$ geschieht. Diese Wahrscheinlichkeit ist ungefähr $2^{-(1-h(\rho))n}$.

Die Wahrscheinlichkeit, dass α bei w unabhängigen Wiederholungen nicht gefunden wird, ist also höchstens

$$\left(1 - 2^{-(1-h(\rho))n}\right)^w \leq e^{-w \cdot 2^{-(1-h(\rho))n}} .$$

Also werden $w = c \cdot 2^{(1-h(\rho))n}$ Wiederholungen gebraucht, um die Fehlerwahrscheinlichkeit unter e^{-c} zu drücken. Daher ist die Laufzeit, bis auf einen polynomiellen Faktor

$$\left(2^{-(1-h(\rho))n} \cdot k^\rho\right)^n = \left(2\rho^\rho(1-\rho)^{(1-\rho)}k^\rho\right)^n .$$

Indem man die Basis nach ρ differenziert und das Ergebnis nach ρ auflöst, erhält man, dass der Ausdruck für $\rho = \frac{1}{k+1}$ minimal wird. Durch einsetzen und elementare Vereinfachungen erhält man dann die Basis $\frac{2k}{k+1}$. Die Komplexität des probabilistischen Algorithmus ist also $\left(\frac{2k}{k+1}\right)^n$. Diese Basis ist für kleine Werte von k:

k	3	4	5	6	7	8
$\frac{2k}{k+1}$	1.5	1.6	1.66667	1.71429	1.75	1.77778

Derandomisierung mit Überdeckungscodes

Bei der probabilistischen Version des Hammingkugel-Algorithmus sind die einzigen zufälligen Entscheidungen, die getroffen werden, die Wahl der Anfangsbewertungen α_i. Um einen deterministischen Algorithmus zu gewinnen, wird der Begriff des Überdeckungscodes verwendet.

Definition. *Eine Menge* C *von Bewertungen ist ein* Überdeckungscode *vom Radius* r*, falls für jede Bewertung* α *gilt* $\alpha \in H(\gamma, r)$ *für ein* $\gamma \in C$*.*

So ist zum Beispiel die oben betrachtete Menge $\{\alpha_0, \alpha_1\}$ aus den beiden konstanten Bewertungen ein Überdeckungscode vom Radius $n/2$.

Hat man einen Überdeckungscode C vom Radius r zur Verfügung, so hat man den folgenden deterministischen Algorithmus für k-SAT, mit Zeitkomplexität $|C| \cdot k^r$:

```
for each γ ∈ C do
    β := suche(γ, r)
    if β ≠ NIL
        then return β
return NIL
```

Um also einen deterministischen Algorithmus zu erhalten, brauchen wir ein Verfahren, um deterministisch einen Überdeckungscode zu konstruieren.

Das Problem, einen möglichst kleinen solchen Code zu konstruieren, ist ein Spezialfall des bekannten NP-schweren Optimierungsproblems SET COVER:

Gegeben:	Eine Menge S, und eine Familie $F \subseteq 2^S$ von Teilmengen von S
Gesucht:	Eine Teilfamilie $C \subseteq F$, die S überdeckt, d.h. $S = \bigcup C$.
Ziel:	Minimiere $\|C\|$.

In unserem Fall ist S die Menge 2^n möglichen Bewertungen, und F die Menge der Hammingkugeln vom Radius r.

Der folgende Greedy-Algorithmus löst das Problem Set Cover: Ein Element $s \in S$ heißt unüberdeckt, falls $s \notin \bigcup C$ ist.

Am Anfang wird $C = \emptyset$ gesetzt. Solange es noch unüberdeckte Elemente in S gibt, wähle ein $c \in F$, dass maximal viele unüberdeckte Elemente enthält, und füge es zi C hinzu.

Falls am Ende der Schleife $\bigcup C = S$ ist, ist C eine Lösung, andernfalls gibt es keine, d.h. $S \setminus \bigcup F \neq \emptyset$.

Es ist eine bekannte Tatsache, dass dieser Greedy-Algorithmus ein $\log|S|$-Approximationsverfahren für Set Cover ist, d.h. er liefert stets eine Lösung C mit $|C| \leq (\log|S|)\cdot|C^*|$, wobei C^* die optimale Lösung, also eine kleinstmögliche Überdeckung, ist.

In unserem Fall ist $\log|S| = O(n)$, also liefert der Greedy-Algorithmus einen Überdeckungscode, der nur um einen Faktor n größer ist als der minimale. Um die Größe, und damit die Komplexität des obigen deterministischen Verfahrens, abschätzen zu können, brauchen wir noch eine obere Schranke an die Größe eines minimalen Überdeckungscodes.

Diese liefert das folgende Lemma. Aus Kardinalitätsgründen muss jeder Überdeckungscode C die Größe $|C| \geq 2^n/V(r)$ haben. Man kann beweisen, dass immer ein Code existiert, der nur um einen Faktor n größer ist.

Lemma 34. *Für alle* n *und* $r = \rho n$ *mit* $0 < \rho \leq 1/2$ *gibt es einen Überdeckungscode vom Radius* r *der Größe*

$$s := n \cdot \sqrt{8n\rho(1-\rho)} \cdot 2^{(1-h(\rho))n} .$$

Beweis. Die Aussage wird mit der probabilistischen Methode bewiesen: es wird gezeigt, dass eine zufällig gewählte Menge von s Bewertungen mit nicht verschwindender Wahrscheinlichkeit ein Überdeckungscode ist. Daher muss es insbesondere einen solchen geben.

Sei also β eine feste Bewertung, und α eine zufällig gewählte. Dann ist $\beta \in H(\alpha, r)$ mit Wahrscheinlichkeit $V(r)/2^n$.

Ist also C eine Menge von s unabhängig zufällig gewählten Bewertungen, so ist die Wahrscheinlichkeit, dass β von C nicht überdeckt wird, höchstens

$$(1 - \frac{V(r)}{2^n})^s \leq e^{-s\frac{V(r)}{2^n}} \leq e^{-s\frac{2^{(h(\rho)-1)n}}{\sqrt{8n\rho(1-\rho)}}} \leq e^{-n} .$$

Somit ist die Wahrscheinlichkeit, dass irgendeine Bewertung nicht überdeckt wird, höchstens $2^n \cdot e^{-}n < 1$, also ist C mit positiver Wahrscheinlichkeit ein Überdeckungscode. □

Da es stets einen Überdeckungscode dieser Größe s gibt, ist insbesondere der optimale Code C^* von der Größe $|C^*| \leq s$, und damit liefert der Greedy-Algorithmus einen Überdeckungscode der Größe

$$O(n) \cdot s = O(n^{5/2}) \cdot 2^{(1-h(\rho))n} .$$

Allerdings ist die Laufzeit des Greedy-Algorithmus für unsere Zwecke zu groß: selbst im günstigsten Fall braucht er mindestens $2^{(1-h(\rho))n}$ Auswahlen, und bei jeder Auswahl muss die Zahl der unüberdeckten Bewertungen für

alle 2^n Hammingkugeln aktualisiert werden, die Laufzeit ist also selbst bei günstigster Implementierung größer als 2^n. Andererseits läuft der Greedy-Algorithmus aber auch, selbst in der einfachsten möpglichen Implementierung, in Zeit 2^{3n}.

Die Lösung besteht in der folgenden einfachen Beobachtung: ist C ein Überdeckungscode vom Radius r für n Variablen, so gewinnt man daraus einen ebensolchen für dn Variablen, indem man die Variablen in d Blöcke mit je n Variablen partitioniert, und alle Bewertungen betrachtet, in denen die Variablen in den Blöcken unabhängig mit allen Bewertungen aus C bewertet werden. Dieser Code C^d hat die Größe $|C|^d$ und den Radius $r \cdot d$.

Es wird also der Greedy-Algorithmus benutzt, um einen Überdeckungscode C für $n/6$ Variablen von Radius $\rho n/6$ zu finden. Dies braucht die Zeit $2^{3n/6} = \sqrt{2}^n$. Dann benutzt man den Code C^6, der die Größe

$$\left[O\left(\left(\frac{n}{6}\right)^{5/2}\right) \cdot 2^{(1-h(\rho))\frac{n}{6}}\right]^6 = O(n^{15}) \cdot 2^{(1-h(\rho))n}$$

hat. Die Gesamtlaufzeit des deterministischen Algorithmus setzt sich also zusammen aus der Zeit $\sqrt{2}^n$ für den Greedy-Algorithmus zum Auffinden des Überdeckungscodes C, plus der Zeit $\left(2^{1-h(\rho)}k^\rho\right)^n$ zum Durchsuchen der Hammingkugeln vom Radius ρn um die Bewertungen im Code C^6. Der deterministische Algorithmus hat also im wesentlichen die gleiche Zeitkomplexität wie der probabilistische Hammingkugel-Algorithmus, nämlich $\left(\frac{2k}{k+1}\right)^n$ für k-SAT.

Der Nachteil dieser Methode ist, dass der Greedy-Algorithmus zur Konstruktion des Überdeckungscodes exponentiell viel Speicherplatz braucht. Es gibt eine zweite Methode, die ohne zusätzlichen Speicherbedarf auskommt, aber dafür eine etwas schlechtere Zeitkomplexität liefert.

Für eine gegebenes $\delta > 0$ wähle ein kleinstmögliches b so dass folgende Ungleichung gilt:

$$b\sqrt{8b\rho(1-\rho)}2^{(1-h(\rho))b} \leq 2^{(1-h(\rho)+\delta)b}$$

Dann existiert nach Lemma 34 ein Überdeckungscode C_b für b Variablen vom Radius ρb der Größe $2^{(1-h(\rho)+\delta)b}$. Dieser endliche Code, den man etwa durch einmalige exhaustive Suche findet, wird im Algorithmus festverdrahtet. Man gewinnt daraus den Code $C_b^{n/b}$ für n Variablen vom Radius ρn, dieser hat die Größe $2^{(1-h(\rho)+\delta)n}$.

Es sei nun ein $\epsilon > 0$ vorgegeben. Dann setzt man $\delta := \log\left(\frac{k+1}{2k}\epsilon + 1\right)$, und benutzt den Code C_b für das zu diesem δ gehörige b wie oben. Die Laufzeit des deterministischen Algorithmus berechnet sich nun wie oben als $\left(2^{1-h(\rho)+\delta} \cdot k^\rho\right)^n$, und mit der Setzung $\rho := 1/(k+1)$ erhält man, dass sie

Basis sich errechnet zu

$$\begin{aligned}\left(2^{1-h(\rho)+\delta}\cdot k^{\rho}\right) &= \left(2^{1+\frac{1}{k+1}\log(\frac{1}{k+1})+\frac{k}{k+1}\log(\frac{k}{k+1})+\delta}\cdot k^{\frac{1}{k+1}}\right)\\ &= \left(2^{1-\frac{1}{k+1}\log(k+1)+\frac{k}{k+1}\log k-\frac{k}{k+1}\log(k+1)+\delta}\cdot 2^{\frac{1}{k+1}\log k}\right)\\ &= 2^{1+\log k-\log(k+1)+\delta} \;=\; 2^{1+\log(\frac{k}{k+1})+\delta}\\ &= 2^{\frac{k}{k+1}\left(\frac{k+1}{2k}\epsilon+1\right)} \;=\; \tfrac{2k}{k+1}+\epsilon\end{aligned}$$

Also erhalten wir als Komplexität den Wert $\left(\frac{2k}{k+1}+\epsilon\right)^n$, für beliebig kleines $\epsilon > 0$.

Für k-SAT mit $k \geq 4$ ist dies der beste bekannte deterministische Algorithmus. Für 3-SAT ist die Komplexität $(1.5+\epsilon)^n$, die von dem oben erwähnten Algorithmus von Schiermeyer unterboten wird. Es ist aber möglich, die Prozedur suche(α, r) im Falle von 3-SAT geschickter zu implementieren, so dass sie statt 3^r nur eine Laufzeit von 2.848^r benötigt. Durch Verwendung dieser verbesserten Prozedur erreicht der deterministische Hammingkugel-Algorithmus die Komplexität 1.481^n, und war damit zur Zeit seiner Veröffentlichung der beste bekannte deterministische Algorithmus für 3-SAT.

6.2 Zufallsirrfahrt

Ein anderer Ansatz besteht darin, nicht nur die Anfangsbewertungen zufälig zu wählen, sondern auch die Verbesserungsschritte. Die innere Schleife führt also eine zufällige Irrfahrt (random walk) im Raum der möglichen Bewertungen durch. Der folgende Algorithmus setzt diese Idee um:

```
repeat w times
    wähle α zufällig
    repeat 3n times
        if α ⊨ F
            then return α
        wähle C = a_1 ∨ ... ∨ a_k mit Cα = 0
        wähle i zufällig aus {1, ..., k}
        α := α with [a_i ← 1]
return NIL
```

Dabei werden die Anfangsbewertungen α wiededrum gleichverteilt und unabhängig gewählt.

Nun ist die Anzahl der Wiederholungen w so zu wählen, dass die Fehlerwahrscheinlichkeit klein wird. Da der Algorithmus nur dann einen Fehler machen, wenn F erfüllbar ist, sei wieder α^* eine fest gewählte Bewertung mit $\alpha^* \models F$.

Wir berechnen die Wahrscheinlichkeit $p(n)$ dafür, dass ein Durchlauf der inneren Schleife α^* findet. Dann ergibt sich die Fehlerwahrscheinlichkeit als $(1-p(n))^w \leq e^{-wp(n)}$, und somit reichen $w = c\frac{1}{p(n)}$ Wiederholungen aus, damit die Fehlerwahrscheinlichkeit höchstens e^{-c} beträgt.

Da $\alpha^* \models F$, gibt es in jeder Klausel $C = (a_1 \vee \ldots \vee a_k)$ in F mindestens ein Literal a_i mit $\alpha^*(a_i) = 1$. Für jede Klausel C wählen wir nun ein festes Literal a_C mit $\alpha^*(a_C) = 1$.

Zur Analyse der Erfolgswahrscheinlichkeit $p(n)$ eines Schleifendurchlaufs wird nun eine Zufallsvariable Z definiert. Bei der Wahl der Anfangsbewertung α ist Z die Hamming-Distanz zwischen α und α^*, also

$$Z := d(\alpha, \alpha^*) \; .$$

Wird in einem Durchlauf der Schleife das Literal a_i aus der Klausel C gewählt, so wird gesetzt

$$Z := \begin{cases} Z-1 & \text{falls } a_i = a_C \\ Z+1 & \text{sonst.} \end{cases}$$

Wird $a_i = a_C$ gewählt, so verringert sich die Distanz $d(\alpha, \alpha^*)$ um 1, da $\alpha^*(a_C) = 1$ ist. Wird dagegen ein Literal $a_i \neq a_C$ gewählt, so wird die Distanz $d(\alpha, \alpha^*)$ entweder um 1 kleiner (falls $\alpha^*(a_i) = 1$ ist) oder um 1 größer (falls $\alpha^*(a_i) = 0$ ist). Es gilt also stets $Z \geq d(\alpha, \alpha^*)$.

Daher ist die Wahrscheinlichkeit, dass α^* gefunden wird, mindestens so gross wie die Wahrscheinlichkeit, dass $Z = 0$ erreicht wird. Diese Wahrscheinlichkeit ist aber wesentlich leichter auszurechnen, als direkt die Wahrscheinlichkeit zu berechnen, dass $\alpha = \alpha^*$ wird, da sich die Zufallsvariable Z sehr einfach verhält.

Man kann sich Z als den Zustand in einem probabilistischen Automaten (einer sog. Markov-Kette), wie in der Abbildung dargestellt, vorstellen. Der Automat startet in einem zufällig zwischen $Z = 0$ und $Z = n$ gewählten Anfangszustand, und geht dann vom einem Zustand Z stets mit Wahrscheinlichkeit $\frac{1}{k}$ in den Zustand $Z-1$ und mit Wahrscheinlichkeit $1-\frac{1}{k}$ in den Zustand $Z+1$ über.

Der Wahrscheinlichkeitsverteilung des Anfangszustand lässt sich nun leicht ausrechnen, da α gleichverteilt gewählt wird, und am Anfang $Z = d(\alpha, \alpha^*)$ ist. Es gibt $\binom{n}{j}$ Bewertungen, die von α^* die Distanz j haben, und jede davon wird mit Wahrscheinlichkeit 2^{-n} gewählt. Der Automat startet daher mit Wahrscheinlichkeit $2^{-n}\binom{n}{j}$ im Zustand $Z = j$.

Es bezeichne nun p_j die Wahrscheinlichkeit, dass der Zustand $Z = 0$ ausgehend von Anfangszustand $Z = j$ erreicht wird. Mit Hilfe der Theorie der

Markov-Ketten errechnet man den Wert $p_j = \left(\frac{1}{k-1}\right)^j$. Daher errechnet sich die Wahrscheinlichkeit, dass $Z = 0$ erreicht wird, zu

$$\begin{aligned}
&\sum_{j=0}^{n} 2^{-n}\binom{n}{j}p_j \\
&= 2^{-n}\sum_{j=0}^{n}\binom{n}{j}\left(\tfrac{1}{k-1}\right)^j \\
&= 2^{-n}\left(1+\tfrac{1}{k-1}\right)^n \\
&= \left(\frac{k}{2(k-1)}\right)^n
\end{aligned}$$

Damit ist auch die Erfolgswahrscheinlichkeit eines Schleifendurchlaufs mindestens $\left(\frac{k}{2(k-1)}\right)^n$, und wir erhalten als obere Schranke an die nötige Zahl von Wiederholungen, und damit die Laufzeit des Algorithmus $O\left(\left(\frac{2(k-1)}{k}\right)^n\right)$. Für $k = 3$ ist dies $(4/3)^n$, und für größere Werte von k erhalten wir die folgenden Basen:

k	3	4	5	6	7	8
$\frac{2(k-1)}{k}$	1.33334	1.5	1.6	1.66667	1.71429	1.75

Verbesserung durch unabhängige Klauseln

Für 3-SAT läßt sich die Komplexität des Irrfahrt-Algorithmus noch leicht verbessern.

Definition. *Eine Menge* $G \subseteq F$ *von Klauseln in* F *heißt* unabhängig, *wenn je zwei Klauseln* C, D *in* F *variablendisjunkt sind, also* $V(C) \cap V(D) = \emptyset$ *ist.*

Ist G unabhängig, so ist $|V(G)| = 3|G|$.

Der Algorithmus berechnet zunächst in Greedy-Manier eine maximale unabhängige Teilmenge G in F wie folgt:

$G := \emptyset$
while es gibt C in F mit $V(C) \cap V(G) = \emptyset$
    $G := G \cup \{C\}$

Nun wird je nach der Größe von G unterschiedlich verfahren:

Da G maximal unabhängig ist, enthält jede der übrigen Klauseln in $F \setminus G$ mindestens eine der Variablen in $V(G)$ enthalten. Wäre nämlich C eine

Klausel, für die dies nicht der Fall ist, so könnte G um diese Klausel erweitert werden.

Ist nun α eine Bewertung der Variablen in $V(G)$, so ist $F\alpha$ eine 2-KNF-Formel, für die das Erfüllbarkeitsproblem in polynomieller Zeit gelöst werden kann. Es gibt insgesamt $7^{|G|}$ Bewertungen, die die Klauseln in G erfüllen.

Ist nun G klein, nämlich $|G| \leq \delta n$ für einen noch zu bestimmenden Parameter $0 < \delta \leq \frac{1}{3}$, so wird für jede der $7^{|G|}$ Bewertungen $\alpha \models G$ das Erfüllbarkeitsproblem für die 2-SAT Instanz $F\alpha$ gelöst. Die Zeitkomplexität ist damit im wesentlichen $7^{\delta n}$.

Ist dagegen $|G| > \delta n$, so wird der Irrfahrt-Algorithmus durchgeführt, wobei allerdings die Anfangsbewertungen nicht gleichverteilt, sondern gemäß der Verteilung $A_G(p_1, p_2, p_3)$ gewählt werden.

Für Wahrscheinlichkeiten $p_1, p_2, p_3 > 0$ mit $3p_1 + 3p_2 + p_3 = 1$ ist die Verteilung $V_G(p_1, p_2, p_3)$ von Bewertungen definiert wie folgt:

Zunächst werden die Variablen in $V(G)$ bewertert, indem für jede Klausel $C = (a_1 \vee a_2 \vee a_3)$ in G Werte $\epsilon_1, \epsilon_2, \epsilon_3$ mit Wahrscheinlichkeit nach der Tabelle

ϵ_1	ϵ_2	ϵ_3		ϵ_1	ϵ_2	ϵ_3	
0	0	0	0	0	1	1	p_2
0	0	1	p_1	1	0	1	p_2
0	1	0	p_1	1	1	0	p_2
1	0	0	p_1	1	1	1	p_3

gewählt werden, und dann die Bewertung

$$[a_1 \leftarrow \epsilon_1, a_1 \leftarrow \epsilon_2, a_1 \leftarrow \epsilon_3]$$

gesetzt wird. Anschliessend wird für jede Variable x in $V(F) \setminus V(G)$ ein Wert $\epsilon \in \{0, 1\}$ zufällig gewählt und $[x \leftarrow \epsilon]$ gesetzt.

Zur Abschätzung der Erfolgswahrscheinlichkeit erinnern wir und, dass die Wahrscheinlichkeit, von einer Anfangsbewertung α aus die erfüllende Bewertung α^* zu erreichen, mindestens

$$\left(\frac{1}{k-1}\right)^{d(\alpha,\alpha^*)}$$

ist. Für $k = 3$ ist dies $\left(\frac{1}{2}\right)^{d(\alpha,\alpha^*)}$.

Es bezeichne $p(\alpha)$ die Wahrscheinlichkeit, dass α als Anfangsbewertung gewählt wird. Dann ist die Erfolgswahrscheinlichkeit

$$p(n) \geq \sum_{\alpha} p(\alpha)\left(\frac{1}{2}\right)^{d(\alpha,\alpha^*)},$$

und der rechte Term ist gerade der Erwartungswert $E[(\frac{1}{2})^{d(\alpha,\alpha^*)}]$ der Zufallsvariable $(\frac{1}{2})^{d(\alpha,\alpha^*)}$.

Wir betrachten nun folgende Zufallsvariablen: X_i sei die Indikatorvariable des Ereignisses, dass $\alpha(x_i) \neq \alpha^*(x_i)$ ist, also

$$X_i = \begin{cases} 1 & \text{falls } \alpha(x_i) \neq \alpha^*(x_i) \\ 0 & \text{sonst.} \end{cases}$$

Dann ist $d(\alpha, \alpha^*) = X_1 + \ldots + X_n$, also ist

$$E\big[(\frac{1}{2})^{d(\alpha,\alpha^*)}\big] = E\big[(\frac{1}{2})^{X_1+\ldots+X_n}\big] \ .$$

Bei gleichverteilter Wahl von α sind nun die Variablen X_i unabhängig, daher können wir ausrechnen

$$E\big[(\frac{1}{2})^{X_1+\ldots+X_n}\big] = \prod_{i=1}^{n} E\big[(\frac{1}{2})^{X_i}\big] \ ,$$

und der Erwartungswert $E[(\frac{1}{2})^{X_i}]$ ist $(\frac{1}{2})^0 \cdot \frac{1}{2} + (\frac{1}{2})^1 \cdot \frac{1}{2} = \frac{3}{4}$, woraus wir wie oben $p(n) \geq (\frac{3}{4})^n$ erhalten.

Im Fall der Wahl von α gemäß $V_G(p_1, p_2, p_3)$ sind die Zufallsvariable X_i nicht mehr unabhängig, aber Abhängigkeit besteht nur zwischen den Variablen innerhalb einer Klausel in G.

Es sei also $G = \{C_1, \ldots, C_{\delta n}\}$ mit $V(C_{i+1}) = \{x_{3i+1}, x_{3i+1}, x_{3i+1}\}$ für $i = 0, \ldots, \delta n - 1$. Wir definieren die Zufallsvariable $S_i = X_{3i+1} + X_{3i+1} + X_{3i+1}$, dann sind diese Variablen S_i wiederum unabhängig, und wir erhalten

$$E\big[(\frac{1}{2})^{X_1+\ldots+X_n}\big] = \prod_{i=1}^{\delta n} E\big[(\frac{1}{2})^{S_i}\big] \cdot \prod_{j=3\delta n+1}^{n} E\big[(\frac{1}{2})^{X_j}\big] \ .$$

Der rechte Faktor berechnet sich wie oben zu $(\frac{3}{4})^{1-3dn}$.

Zur Berechnung des linken Faktors brauchen wir die Erwartungswerte der Variablen $(\frac{1}{2})^{S_i}$. Diese hängen aber davon ab, wieviele der Literale in C_i von α^* erfüllt werden.

Falls α^* genau ein Literal a in C_i mit 1 bewertet, so hat S_j den Wert 0, wenn α auf C_i mit α^* übereinstimmt, also nur a mit 1 und die anderen Literale in C_i mit 0 bewertet. Dies passiert mit Wahrscheinlichkeit p_1.

Weiter hat S_i den Wert 1, wenn α das Literal a und ein weiteres mit 1 bewertet, dafür gibt es 2 Möglichkeiten, die jeweils mit Wahrscheinlichkeit p_2 eintreten.

S_i hat den Wert 2, wenn α alle drei Literale in C_i mit 1 bewertet, was mit Wahrscheinlichkeit p_3 eintritt, oder wenn α das Literal a mit 0 und ein anderes mit 1 bewertet, was mit Wahrscheinlichkeit $2p_1$ eintritt.
Schließlich hat S_i den Wert 3, wenn a von α mit 0 und die zwei anderen Literale mit 1 bewertet werden, was mit Wahrcheinlichkeit p_2 eintritt.
Insgesamt errechnet sich der Erwartungswert in diesem Fall also als

$$E\left[\left(\tfrac{1}{2}\right)^{S_i}\right] = \left(\tfrac{1}{2}\right)^0 p_1 + \left(\tfrac{1}{2}\right)^1 2p_2 + \left(\tfrac{1}{2}\right)^2 (2p_1 + p_3) + \left(\tfrac{1}{2}\right)^3 p_2 = \tfrac{3}{2}p_1 + \tfrac{9}{8}p_2 + \tfrac{1}{4}p_3 \,.$$

Im Fall, dass α^* genau zwei Literale in C_i mit 1 bewertet, errechnet man auf ähnliche Weise

$$E\left[\left(\tfrac{1}{2}\right)^{S_i}\right] = \left(\tfrac{1}{2}\right)^0 p_2 + \left(\tfrac{1}{2}\right)^1 (2p_1 + p_3) + \left(\tfrac{1}{2}\right)^2 2p_2 + \left(\tfrac{1}{2}\right)^3 p_1 = \tfrac{9}{8}p_1 + \tfrac{3}{2}p_2 + \tfrac{1}{2}p_3 \,.$$

Im letzten Fall, wo α^* alle drei Literale in C_i mit 1 bewertet, errechnet man schließlich

$$E\left[\left(\tfrac{1}{2}\right)^{S_i}\right] = \left(\tfrac{1}{2}\right)^0 p_3 + \left(\tfrac{1}{2}\right)^1 3p_2 + \left(\tfrac{1}{2}\right)^2 3p_1 = \tfrac{3}{4}p_1 + \tfrac{3}{2}p_2 + p_3 \,.$$

Für $j = 1, 2, 3$ sei also m_j die Anzahl der Klauseln in G, in denen α^* genau j Literale erfüllt. Dann ist also der linke Faktor des obigen Produktes gleich

$$\left(\tfrac{3}{2}p_1 + \tfrac{9}{8}p_2 + \tfrac{1}{4}p_3\right)^{m_1} \cdot \left(\tfrac{9}{8}p_1 + \tfrac{3}{2}p_2 + \tfrac{1}{2}p_3\right)^{m_2} \cdot \left(\tfrac{3}{4}p_1 + \tfrac{3}{2}p_2 + p_3\right)^{m_3} \,.$$

und durch Lösen des linearen Gleichungssystems aus der Gleichheit der drei Basen zusammen mit der Einschränkung $3p_1 + 3p_2 + p_3 = 1$ erhält man die Werte

$$p_1 = \tfrac{4}{21}, \qquad p_2 = \tfrac{2}{21}, \qquad p_3 = \tfrac{1}{7} \quad ,$$

für die sich die Basis als $\frac{3}{7}$ ergibt. Damit ist die Erfolgswahrscheinlichkeit mindestens

$$p(n) \geq \left(\tfrac{3}{7}\right)^{m_1+m_2+m_3} \cdot \left(\tfrac{3}{4}\right)^{1-3\delta n} = \left(\tfrac{3}{7}\right)^{\delta n} \cdot \left(\tfrac{3}{4}\right)^{1-3\delta n} \,,$$

und damit die nötige Anzahl von Wiederholungen und somit auch die Laufzeit des Algorithmus

$$\left(\left(\tfrac{4}{3}\right)^{1-3\delta} \cdot \left(\tfrac{7}{3}\right)^{\delta}\right)^n \,.$$

Man beobachtet nun, dass dieser Wert in δ monoton fallend ist, während die Laufzeit $7^{\delta n}$ im anderen Zweig monoton in δ ansteigt. Der minimale Wert der Laufzeit ergibt sich also an der Schnitstelle, wo

$$7^{\delta} = \left(\tfrac{4}{3}\right)^{1-3\delta} \cdot \left(\tfrac{7}{3}\right)^{\delta}$$

ist. Dieser Wert von δ ist ungefähr 0.14666, und der Wert 7^{δ} an dieser Stelle ist ungefähr 1.33026, also erhalten wir für den Algorithmus insgesamt eine Laufzeit von $O(1.33026^n)$.
Es gibt noch eine verbesserte Version des Algorithmus, die eine Laufzeit von $O(1.32793^n)$ erreicht, diese wird aber aus Platzgründen hier nicht behandelt.

6.2.1 Derandomisierung von Moser und Scheder

Aus der Analyse des Hammingkugel-Algorithmus erhält man die

Proposition 35. *Ist* A *ein Algorithmus, der in Zeit* $O(t^r)$ *die Hamming-Kugel vom Radius* r *durchsucht, so gewinnt man daraus einen Algorithmus, der* k*-SAT in Zeit* $(2t/(t+1))^n$ *löst. Dieser ist deterministisch, falls* A *deterministisch ist.*

Mit dem beim oben behandelten Hammingkugel-Algorithmus verwendeten Algorithmus suche(α, r) mit der Laufzeit $O(k^r)$ erhält man somit die bekannte Laufzeit $(2k/(k+1))^n$ dieses Algorithmus.

Moser und Scheder geben einen deterministischen Algorithmus an, der für beliebiges $\delta > 0$ die Hamming-Kugel vom Radius r in Zeit $O((k-1+\delta)^r)$ durchsucht. Mit der obigen Proposition erhält man daraus also einen deterministischen Algorithmus, der k-SAT in Zeit $O\left(\left(\frac{2(k-1)}{k}+\epsilon\right)^n\right)$ für ein beliebiges $\epsilon > 0$ löst.

Der Algorithmus verwendet eine Form von Derandomisierung des Verbesserungsschrittes in Schönings Irrfahrt-Algorithmus. Betrachten wir diesen zunächst für den Fall von 3-SAT. Sei α^* eine erfüllende Bewertung mit Hamming-Distanz $d(\alpha, \alpha^*) \leq r$ zur aktuellen Bewertung α.

Sind C und C' unabhängige Klauseln mit $\alpha(C) = \alpha(C') = 0$, und der Algorithmus wählt die Klausel C zur Verbesserung, so ist danach noch immer $\alpha(C') = 0$ und kann als nächstes ausgewählt werden. Also kann der Algorithmus auch so aufgefasst werden, dass für mehrere unabhängige Klauseln $C_1, \ldots, C_t$ mit $\alpha(C_1) = \ldots = \alpha(C_t) = 0$ jeweils gleichzeitig ein Literal $a_i \in C_i$ gewählt und α so geändert wird, dass $\alpha(a_i) = 1$ gesetzt wird.

Werden dabei alle t Literale a_i so gewählt, dass $\alpha^*(a_i) = 1$ ist, verringert sich die Distanz von α zu α^* um t. Dies passiert aber nur mit einer geringen Wahrscheinlichkeit von $1/3^t$. Ein sehr viel wahrscheinlicheres Ergebnis ist, dass $\alpha^*(a_i) = 1$ für $2t/3$ der Literale a_i ist, und $\alpha^*(a_i) = 0$ für die übrigen $t/3$, wodurch sich die Distanz $d(\alpha, \alpha^*)$ um $t/3$ verringert. Eine solche Auswahl kann nun deterministisch getroffen werden.

Die Auswahl je eines Literals aus einer Menge von t Klauseln der Breite k wird durch einen String $w = w_1 w_2 \ldots w_k \in \{1, \ldots, k\}^t$ gegeben. Für zwei solche Strings w, w' ist die Hamming-Distanz $d(w, w') = \left|\{i; w_i \neq w'_i\}\right|$, und die Hamming-Kugel um w vom Radius r ist $H_k(w, r) = \{w';\ d(w, w') \leq r\}$.

Ein k-stelliger Überdeckungscode vom Radius r ist eine Menge $C \subseteq \{1, \ldots, k\}^t$ so dass jeder String $w' \in \{1, \ldots, k\}^t$ in der Hamming-Kugel $H_k(w, r)$ für ein $w \in C$ liegt.

Lemma 36. *Für alle* $r \leq t$ *gibt es einen* k*-stelligen Überdeckungscode* C *vom Radius* r *der Größe*

$$|C| \;\leq\; \frac{t\ln(k)k^t}{\binom{t}{r}(k-1)^r} \;\leq\; t^2(k-1)^{t-2t/k}$$

Der Algorithmus verwendet einen Überdeckungscode $C \subseteq \{1,\ldots,k\}^t$ vom Radius t/k, für eine noch zubestimmende Konstante t. Dieser Code C der Größe $|C| \leq t^2(k-1)^{t-2t/k}$ wird vorab durch exhaustive Suche berechnet und im Algorithmus fest verdrahtet.

Der Algorithmus $MS(\alpha, r)$ soll nun die Hammingkugel $H(\alpha, r)$ durchsuchen, und wir nehmen an $\alpha^* \in H(\alpha, r)$ sei eine feste Bewertung mit $\alpha^* \models F$.

Zunächst wird eine maximale unabhängige Menge $G = \{C_1,\ldots,C_m\}$ von Klauseln mit $\alpha(C_i) = 0$ füs alle $1 \leq i \leq m$ berechnet. Es werden dann zwei Fälle abhängig von der Anzahl m dieser Klauseln unterschieden.

Ist $m < t$, so wird für jede Bewertung $\beta : V(G) \to \{0,1\}$ der Variablen von G der Algorithmus $suche(\alpha, r)$ auf der Formel $F\beta$ aufgerufen.

Da G maximal unabhängig ist, ist für jedes solche β die Formel $F\beta$ in $(k-1)$-KNF, denn jede Klausel in $F \setminus G$ muss mindestens eine Variable aus $V(G)$ enthalten.

Daher benötigt $suche(\alpha, r)$ auf $F\beta$ die Zeit $(k-1)^r$, und damit ist die Laufzeit insgesamt $2^{km}(k-1)^r = O((k-1)^r)$.

Ist dagegen $m \geq t$, so betrachten wir die Klausel $C_1,\ldots,C_t$, und es sei für $1 \leq i \leq t$ die Klausel $C_i = a_{i,1} \vee \ldots \vee a_{i,k}$.

Für einen String $w \in \{1,\ldots,k\}^t$ definiere die Bewertung $\alpha[w]$, die definiert ist wie α, nur dass für $1 \leq i \leq t$ der Wert des Literals a_{i,w_i}, also des Literals in C_i an der durch w gewählten Stelle abgeändert wird zu $1 - \alpha(a_{i,w_i})$.

Sei w^* ein String mit $\alpha^*(a_{i,w_i^*}) = 1$, der in jeder Klausel ein durch α^* erfülltes Literal auswählt. Dann ist die Distanz

$$d(\alpha[w^*], \alpha^*) \leq d(\alpha, \alpha^*) - t \leq r - t\,.$$

Da C ein Überdeckungscode vom Radius t/k ist, gibt es einen String $w \in C$ mit $d(w, w^*) \leq t/k$, also stimmt w mit w^* an mindestens $t - t/k$ Stellen überein, und unterscheidet sich von w^* an höchstens t/k Stellen. Daher ist

$$d(\alpha[w], \alpha^*) \leq d(\alpha, \alpha^*) - (t - t/k) + t/k \leq r - (t - 2t/k)\,.$$

Es sei also $\Delta := (t-2t/k)$. Es wird nun für jedes $w \in C$ rekursiv $MS(\alpha[w], r-\Delta)$ aufgerufen, und wegen der obigen Argumentation ist einer dieser rekursiven Aufrufe erfolgreich.

Es wird also ein Rekursionsbaum aufgespannt mit dem Verzweigungsgrad $|C|$ und der Tiefe r/Δ, dessen Größe — und damit die Laufzeit — ist also beschränkt durch

$$|C|^{r/\Delta} \leq \left(t^2(k-1)^{\Delta}\right)^{r/\Delta} \leq \left((k-1)t^{2/\Delta}\right)^r .$$

Für gegebens $\delta > 0$ ist also t so zu wählen, dass $(k-1)t^{2/\Delta} \leq (k-1)+\delta$ ist, dann ist die Laufzeit wie behauptet $O((k-1+\delta)^r)$.

Der Algorithmus von Moser und Scheder ist der derzeit beste deterministische Algorithmus für k-SAT, für $k \geq 4$. Für 3-SAT gibt es eine derandomisierte Version der verbesserten Version des Irrfahrt-Algorithmus aus dem vorherigen Abschnitt, der auch dessen Laufzeitschranke erreicht.

6.3 Algorithmus von Paturi, Pudlák und Zane

In diesem Abschnitt behandeln wir einen probabilistischen Algorithmus, der der Vorgehensweise bei DPLL ähnlicher ist als die vorher betrachteten.

Wie bei einem DPLL-Verfahren werden die Variablen der Reihe nach betrachtet und ihnen Werte zugewiesen. Dabei werden Variablen in Einerklauseln natürlich so bewertet, dass diese erfüllt sind, die übrigen Variablen werden zufällig bewertet. Die Reihenfolge, in der die Variablen betrachtet werden, ist durch eine Permutation $\pi \in S_n$ der Indizes $1, \ldots, n$ vorgegeben, die ebenfalls zufällig gewählt wird. Ist am Ende die Formel erfüllt, wird die so gefundene Bewertung audsgegeben, andernfalls wird sie verworfen und mit einer neuen Permutation von vorn begonnen.

Die Prozedur suche(π) betrachtet die Variablen in der durch π vorgegebenen Reihenfolge und weist ihnen einen zufälligen Wert zu, außer wenn ein Wert bereits durch Einerklauseln erzwungen ist.

```
suche(π)
  for i := 1 to n do
      if Einerklausel x^ε_π(i) in F
          then ε_i := ε
          else wähle ε_i zufällig aus {0, 1}
      F := F[x_π(i) ← ε_i]
  if α ⊨ F
      then return α
      else return NIL
```

Diese Prozedur wird in einer Schleife eine bestimmte Anzahl w mal wiederholt.

```
  repeat w times
```

```
    wähle π ∈ S_n zufällig
    β := suche(π)
    if β ≠ NIL
        then return β
return NIL
```

Wie bei den probablilistischen Algorithmen im vorhergehenden Kapitel berechnen wir die Wahrscheinlichkeit $p(n)$ dafür, dass ein Schleifendurchlauf erfolgreich ist, also eine erfüllende Bewertung findet, im Fall dass F erfüllbar ist. Daraus ergibt sich, dass $w = O(1/p(n))$ Wiederholungen genügen, um die Fehlerwahrscheinlichkeit beliebig klein zu halten, und damit eine Laufzeit von $O(1/p(n))$.

Definition. *Für eine Menge A von Bewertungen und $\alpha \in A$ ist*

$$N_A(\alpha) := |H(\alpha, 1) \cap A| - 1$$

die Anzahl der Nachbarn von α in A.

Lemma 37. *Ist $A \neq \emptyset$, so gilt*

$$\sum_{\alpha \in A} 2^{-N_A(\alpha)} \geq 1 \,.$$

Beweis. Die Aussage wird durch Induktion nach n bewiesen. Der Induktionsanfang für $n = 0$ ist trivial.

Sei also $n > 0$, und gelte die Aussage für Mengen von Bewertungen von $n-1$ Variablen. Für eine Bewertung α der Variablen $x_1, \ldots, x_{n-1}$ und $i = 0, 1$ sei $\alpha i := \alpha \cup [x_n \leftarrow i]$.

Sei $A \neq \emptyset$ eine Menge von Bewertungen aller n Variablen. Für $i = 0, 1$ sei $A_i := \{\, \alpha \,;\, \alpha i \in A \,\}$. Für mindestens ein i muss $A_i \neq \emptyset$ sein.

Ist $A_0 = \emptyset$, so muss $A_1 \neq \emptyset$ sein. Also ist

$$\sum_{\alpha \in A} 2^{-N_A(\alpha)} \;=\; \sum_{\alpha \in A_1} 2^{-N_A(\alpha 1)} \;=\; \sum_{\alpha \in A_1} 2^{-N_{A_1}(\alpha)} \;\geq\; 1 \,,$$

wobei die letzte Ungleichung nach der Induktionshypothese gilt. Die letzte Gleichheit gilt, da für $\alpha \in A_1$ der einzige Nachbar von $\alpha 1$, der nicht von der Form β_1 ist, gerade $\alpha 0$ ist. $A_0 = \emptyset$ ist aber $\alpha 0 \notin A$.

Der Fall $A_1 = \emptyset$ ist symmetrisch. Seien nun A_0 und A_1 beide nichtleer. Dann ist

$$\sum_{\alpha \in A} 2^{-N_A(\alpha)} \;=\; \sum_{\alpha \in A_0} 2^{-N_A(\alpha 0)} + \sum_{\alpha \in A_1} 2^{-N_A(\alpha 1)}$$

$$\geq \sum_{\alpha\in A_0} 2^{-N_{A_0}(\alpha)-1} + \sum_{\alpha\in A_1} 2^{-N_{A_1}(\alpha)-1}$$
$$\geq \frac{1}{2} + \frac{1}{2} = 1 .$$

Dabei gilt die vorletzte Ungleichung wiederum, da es für $i = 0, 1$ höchstens einen Nachbarn von αi gibt, der nicht von der Form β_i ist, nämlich $\alpha 0$ $\square$

Zur Berechnung der Erfolgswahrscheinlichkeit $p(n)$ sei $A := \{\, \alpha \,;\, \alpha \models F \,\}$ die Menge der F erfüllenden Bewertungen, und sei $\alpha \in A$. Definiere

$$j(\alpha) := n - N_A(\alpha)$$

und für $i = 1, \ldots, n$ sei $\alpha_i := \alpha$ with $[x_i \leftarrow 1 - \alpha(x)]$. Das heißt, es gibt $j(\alpha)$ Variablen x_i mit $\alpha_i \not\models F$, diese heißen die *kritischen* Variablen.

Für jede kritische Variable x_i gibt es eine *kritische* Klausel C_i in F, derart dass das einzige Literal a in C_i mit $\alpha(a) = 1$ entweder x_i oder $\neg x_i$ ist.

Für eine Permutation $\pi \in S_n$ sei

$$R(\pi, \alpha) = \left\{ x_i \text{ kritisch} \,;\, \pi^{-1}(i) \geq \pi^{-1}(j) \text{ für alle } x_j \in V(C_i) \right\}$$

die Menge derjenigen kritischen Variablen x_i, die unter den Variablen in ihrer kritischen Klausel C_i in der Permutation π die letzte sind, und $r(\pi, \alpha) = |R(\pi, \alpha)|$.

Da die Permutation π gleichverteilt gewählt wird, ist jede der k Variablen in C_i mit gleicher Wahrscheinlichkeit $1/k$ die letzte in π. Jede der $j(\alpha)$ kritischen Variablen ist also mit Wahrscheinlichkeit $1/k$ in $R(\pi, \alpha)$, daher ist der erwartete Wert $E[r(\pi, \alpha)] = j(\alpha)/k$.

Mit jeder Variablen $x_i \in R(\pi, \alpha)$ erhöht sich die Wahrscheinlichkeit dafür, dass α bei suche(π) gefunden wird: wenn die anderen Variablen y in C_i zufällig den richtigen Wert $\alpha(y)$ erhalten, dann erhält auch x_i den Wert $\alpha(x_i)$, da zu dem Zeitpunkt, wo der Algorithmus x_i betrachtet, die Klausel C_i Einerklausel ist.

Daher ist die Wahrscheinlichkeit $p_{\alpha,\pi}$ dafür dass α gefunden wird, falls die Permutation π gewählt wurde, mindestens $p_{\alpha,\pi} \geq 2^{-n+r(\pi,\alpha)}$.

Also ist die Wahrscheinlichkeit p_α dafür, dass α gefunden wird, insgesamt

$$p_\alpha = \sum_{\pi\in S_n} \frac{1}{n!} p_{\alpha,\pi} = \sum_{\pi\in S_n} \frac{1}{n!} 2^{-n+r(\pi,\alpha)}$$
$$= 2^{-n} \sum_{\pi\in S_n} \frac{1}{n!} 2^{r(\pi,\alpha)} = 2^{-n} \cdot E[2^{r(\pi,\alpha)}] .$$

Der Erwartungswert $E[2^{r(\pi,\alpha)}]$ ist aber nach der Jensen'schen Ungleichung mindesten $2^{E[r(\pi,\alpha)]}$, also ist

$$p_\alpha \;\geq\; 2^{-n+E[r(\pi,\alpha)]} \;\geq\; 2^{-n+j(\alpha)/k}\,.$$

Die Jensen'sche Ungleichung besagt, dass für eine Zufallsvariable X, die jeden der Werte $a_1,\ldots,a_n$ mit Wahrscheinlichkeit $1/n$ annimmt, gilt $E[2^X]\cdot 2^{E[x]}$. Dies rechnet man nach, indem man den Logarithmus zur Basis 2 betrachtet, also zeigt $\log E[2^X]\geq E[X]$. Es ist nämlich

$$\begin{aligned}\log E[2^X] \;&=\; \log\Big(\frac{1}{n}\sum_{i=1}^{n}2^{a_i}\Big) \;\geq\; \log\Big(\big(\prod_{i=1}^{n}2^{a_i}\big)^{1/n}\Big)\\ &=\; \frac{1}{n}\log\prod_{i=1}^{n}2^{a_i} \;=\; \frac{1}{n}\sum_{i=1}^{n}\log 2_i^a \;=\; \frac{1}{n}\sum_{i=1}^{n}a_i \;=\; E[X]\end{aligned}$$

wobei die erste Ungleichung nach dem Satz vom arithmetischen und geometrischen Mittel gilt.

Die gesamte Erfolgswahrscheinlichkeit lässt sich nun beschränken durch

$$\begin{aligned}p(n) \;&=\; \sum_{\alpha\in A}p_\alpha \;\geq\; \sum_{\alpha\in A}2^{-n+j(\alpha)/k}\\ &=\; \sum_{\alpha\in A}2^{-n+n/k+(j(\alpha)-n)/k} \;=\; 2^{-n+n/k}\sum_{\alpha\in A}2^{-N(\alpha)/k}\\ &\geq\; 2^{-n+n/k}\sum_{\alpha\in A}2^{-N(\alpha)} \;\geq\; 2^{-n+n/k}\,.\end{aligned}$$

wobei die letzte Ungleichung nach Lemma 37 gilt.

Also ist $p(n)\geq 2^{-n+n/k}$, und somit erhalten wir als Schranke an die Laufzeit $O\big((2^{1-1/k})^n\big)$. Die Basis errechnet sich für kleine Werte von k wie folgt:

k	3	4	5	6	7	8
$2^{1-1/k}$	1.58741	1.68180	1.74111	1.78180	1.81145	1.83401

Dieser Algorithmus war zum Zeitpunkt seiner Veröffentlichung 1997 der erste, der eine bessere Komplexität für k-SAT mit $k\geq 4$ erreicht als der von Monien und Speckenmeyer. Ausserdem war dies der erste theoretisch analysierte probabilistische Algorithmus für k-SAT.

Eine verbesserte Form dieses Algorithmus wurde von Paturi, Pudlák, Saks und Zane etwas später angegeben. Der Algorithmus arbeitet genau wie der hier vorgestellte, aber es wird ein Präprozessor vorgeschaltet, der die Formel um alle Klauseln ergänzt, die aus den Eingabeklauseln durch Resolution erzeugt werden können, und deren Größe durch einen (von k und n

abhängigen) Parameter s beschränkt ist. Die Analyse ist zu komplex, um hier präsentiert zu werden.

Dieser verbesserte Algorithmus ist derzeit der beste probabilistische Algorithmus für k-SAT mit $k \geq 5$, er erreicht die Komplexität $O(b_k^n)$ für die folgenden Werte der Basis b_k:

k	3	4	5	6	7	8
b_k	1.36406	1.49579	1.56943	1.63788	1.68753	1.72521

Schon seit seiner Veröffentlichung war bekannt, dass sich für den Algorithmus von Paturi, Pudlák, Saks und Zane eine schärfere Schranke für die Erfolgswahrscheinlichkeit bei Formeln in 3-KNF und 4-KNF zeigen lässt, die genau eine erfüllende Bewertung haben. Kürzlich wurde von Timon Hertli gezeigt, dass diese Analyse auch für allgemeine Formeln in 3-KNF und 4-KNF gilt. Dies führt zu einer verbesserten Laufzeit des Algorithmus von $O(1.308^n)$ für 3-SAT und $O(1.469^n)$ für 4-SAT, womit dieser auch für diese Fälle der beste bekannte probabilistische Algorithmus ist.