Erinnerung: Graphenprobleme in NP

Problem INDEPENDENT SET

Instanz: Graph $G = (V, E)$, $k \in \mathbb{N}$
Frage: Gibt es unabhängige Menge I mit $|I| \geq k$?

unabhängige Menge: $I \subseteq V$ mit $\{x, y\} \notin E$ für alle $x, y \in I$

Problem VERTEX COVER

Instanz: Graph $G = (V, E)$, $k \in \mathbb{N}$
Frage: Gibt es vertex cover U mit $|U| \leq k$?

Vertex cover: $U \subseteq V$ mit $e \cap U \neq \emptyset$ für alle $e \in E$

NP-vollständige Graphenprobleme

Theorem

INDEPENDENT SET *ist NP-vollständig.*

Beweis: Durch Reduktion von 3-SAT.

Korollar

VERTEX COVER ist NP-vollständig.

Beweis: INDEPENDENT SET $\leq_P$ VERTEX COVER

Reduktion von INDEPENDENT SET auf VERTEX COVER

Lemma

INDEPENDENT SET $\leq_P$ VERTEX COVER

Beweis: $I \subseteq V$ ist unabhängig gdw. $V \setminus I$ ein Vertex Cover ist.

Also hat G eine unabhängige Menge der Größe k
gdw. G eine Vertex Cover der Größe $|V| - k$ hat.

$\rightsquigarrow$ $(G, k) \mapsto (G, |V| - k)$ ist polynomielle Reduktion von INDEPENDENT SET auf VERTEX COVER.

Ist $I \subseteq V$ unabhängig, dann ist $V \setminus I$ ein Vertex Cover: da es keine Kanten zwischen zwei Knoten in I gibt, muss jede Kante mindestens ein Ende haben, das nicht in I, also in $V \setminus I$ liegt.

Die umgekehrte Richtung gilt ganz analog: ist $C \subseteq V$ ein Vertex Cover, so ist $V \setminus C$ unabhängig: jede Kante hat ein Ende in C, also gibt es keine Kante zwischen zwei Knoten in $V \setminus C$.

Reduktion von 3-SAT auf INDEPENDENT SET

Sei $F = C_1 \wedge \ldots \wedge C_m$, und $C_i = a_{i,1} \vee a_{i,2} \vee a_{i,3}$

Konstruiere Graphen $G(F)$ mit:

$G(F)$ hat unabh. Menge I mit $|I| \geq m$ gdw. F erfüllbar ist.

- Knoten: $v_{i,j}$ für jedes $1 \leq i \leq m$ und $j = 1, 2, 3$
- Dreieck $v_{i,1}$, $v_{i,2}$, $v_{i,3}$ für jedes i
- Kanten $\{v_{i,j}, v_{i',j'}\}$ wenn $a_{i,j} = \neg a_{i',j'}$

Es gilt:

- I mit $|I| = m$ $\rightsquigarrow$ α_I so dass $\alpha_I \models F$
- α mit $\alpha \models F$ $\rightsquigarrow$ $I(\alpha)$ mit $|I(\alpha)| = m$

Der Graph $G(F)$ hat also für jede Klausel in F ein Dreieck, in dem jeder Knoten einem der Literale in der Klausel entspricht. Eine unabhängige Menge kann also höchstens einen Knoten aus jedem dieser Dreiecke enthalten, und somit höchstens die Größe m haben.

Zusätzlich gibt es Kanten zwischen Knoten, die komplementären Literalen entsprechen, also x und $\bar{x}$ für eine Variable x. Eine unabhängige Menge kann also höchstens einen dieser Knoten enthalten.

Ist also I unabhängig mit $|I| = m$, dann muss I aus jedem Dreieck genau einen Knoten enthalten. Wir erhalten daraus α_I, indem wir die entsprechenden Literale auf 1 setzen. Die Kanten vom zweiten Typ stellen sicher, dass dies wirklich eine gültige Bewertung ist (also nicht x und $\bar{x}$ auf 1 gesetzt werden).

Umgekehrt setzt eine Bewertung $\alpha \models F$ in jeder Klausel ein Literal zu 1, und wir erhalten daraus I_α, indem wir aus jedem Dreieck einen Knoten wählen, der einem Literal mit Wert 1 entspricht. Da α eine Bewertung ist, ist diese Menge tatsächlich unabhängig, und sie enthält offensichtlich m Knoten.

Hamilton-Kreise

Ein Hamilton-Kreis in $G = (V, E)$ mit $|V| = n$ ist ein geschlossener Weg

$$v_1, \ldots, v_n, v_{n+1} = v_1 \quad \text{mit } \{v_i, v_{i+1}\} \in E \text{ für alle } i \leq n$$

in dem jeder Knoten genau einmal vorkommt, also $\{v_1, \ldots, v_n\} = V$

Problem HAMILTON-KREIS

Instanz:	Graph $G = (V, E)$
Frage:	Gibt es einen Hamilton-Kreis in G ?

Problem GERICHTETER HAMILTON-KREIS

Instanz:	gerichteter Graph $G = (V, E)$ mit $E \subseteq V \times V$
Frage:	Gibt es einen Hamilton-Kreis in G ?

Ein Hamilton-Kreis ist also ein geschlossener Weg duch den Graphen, der jeden Knoten genau einmal besucht.

Ein ähnlicher Begriff ist der Euler-Kreis: dies ist ein geschlosssener Weg, der jede Kante genau einmal durchläuft (wie beim bekannten Spiel *Das Haus vom Nikolaus*). Eine offensichtliche notwendige Bedingung für die Existenz eines Euler-Kreises ist, dass an jedem Knoten gerade viele Kanten inzidieren. Diese Bedingung ist auch hinreichend: ein Graph hat genau dann einen Euler-Kreis, wenn jeder Knoten geraden Grad hat, also gerade viele Kanten mit ihm inzidieren. Daher ist das Problem, ob ein Euler-Kreis existiert, in P. Das Problem Hamilton-Kreis ist schwieriger, weil man keine solche äquivalente einfach zu überprüfende Bedingung kennt.

NP-Vollständigkeit von Hamilton-Kreis

Satz

Hamilton-Kreis *ist NP-vollständig.*

Lemma

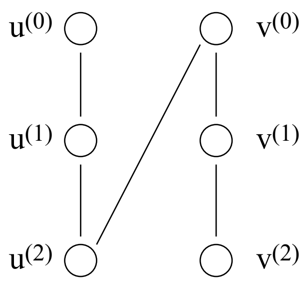
3-*SAT* $\leq_P$ Gerichteter Hamilton-Kreis

Lemma

Gerichteter Hamilton-Kreis $\leq_P$ Hamilton-Kreis

Zuerst zeigen wir die einfachere Reduktion von Gerichteter Hamilton-Kreis auf Hamilton-Kreis. Es muss also aus einem gegebenen gerichteten Graphen G ein ungerichteter Graph G_U konstruiert werden, der genau dann einen Hamilton-Kreis hat, wenn G einen gerichteten Hamilton-Kreis hat.

Für jeden Knoten v in G hat G_U drei Knoten $v^{(0)}$, $v^{(1)}$ und $v^{(2)}$, und Kanten zwischen $v^{(0)}$und $v^{(1)}$ sowie zwischen $v^{(1)}$ und $v^{(2)}$. Ist (u, v) eine Kante in G, so gibt es in G_U eine Kante zwischen $u^{(2)}$ und $v^{(0)}$, wie im Bild dargestellt.



Ist nun $v_1, v_2, \ldots, v_n, v_1$ ein gerichteter Hamilton-Kreis in G, so ist

$$v_1^{(0)}, v_1^{(1)}, v_1^{(2)}, v_2^{(0)}, v_2^{(1)}, v_2^{(2)}, \ldots, v_n^{(0)},, v_n^{(1)}, v_n^{(2)}, v_1^{(0)}$$

ein Hamilton-Kreis in G_U.

Für die Umkehrung ist zu beobachten, dass jeder Hamilton-Kreis in G_U entweder alle Knotentripel in der Reihenfolge $v_1^{(0)}, v_1^{(1)}, v_1^{(2)}$ oder alle in der umgekehrten Reihenfolge durchlaufen muss. Wird die Richtung an einer Stelle gewechselt, also z.B.

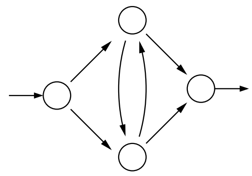
$$\ldots u^{(0)}, u^{(1)}, u^{(2)}, v^{(0)}, w^{(2)}, w^{(1)}, w^{(0)} \ldots$$

dann kann der Knoten $v^{(1)}$ nicht mehr erreicht werden.

Dann kann aber aus einem Hamilton-Kreis in G_U entweder ein gerichteter Hamilton-Kreis in G, oder einer in umgekehrter Richtung abgelesen werden.

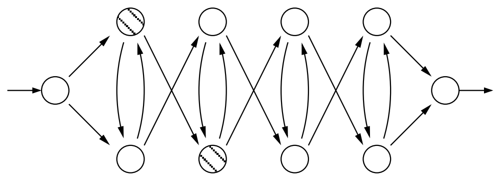
Sei $F = C_1 \wedge \ldots \wedge C_m$ eine Formel in 3-KNF, in den Variablen $x_1, \ldots, x_n$. Wir konstruieren daraus einen gerichteten Graphen, der genau dann einen Hamilton-Kreis hat, wenn F erfüllbar ist. Zuerst brauchen wir eine Repräsentation von Variablen und ihrer Bewertungen.

Dazu könnte der rechts abgebildete Teilgraph dienen, da ihn ein Hamilton-Kreis auf genau zwei Weisen durchlaufen kann: von links erst nach oben, dann nach unten und dann nach rechts oder umgekehrt.



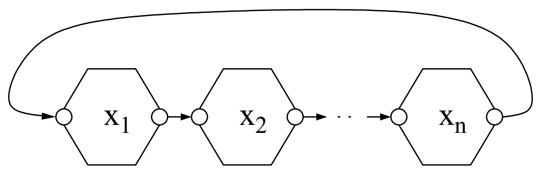
Eine Repräsentation der Klauseln könnte dann mit den senkrechten Kanten verbunden werden, die positiven Vorkommen mit der abwärts und die negativen mit der aufwärts verlaufenden, so dass der Kreis die erfüllten Literale durchläuft.

Da eine Variable mehrfach vorkommen kann, wird tatsächlich für jede Variable ein größerer Teilgraph konstruiert, der immer noch genau zwei Arten des Durchlaufs erlaubt.



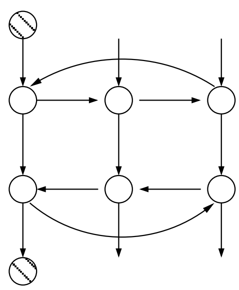
Kommt die Variable x_i in m_i Klauseln vor, so hat der obige Teilgraph $2m_i + 4$ Knoten, und die Teilgraphen, die Klauseln repräsentieren, werden mit den schräg verlaufenden Kanten verbunden, z.B. das erste positive Vorkommen von x_i mit der Kante zwischen den schraffierten Knoten.

Diese Teilgraphen für die n Variablen werden in einem großen Kreis miteinander verbunden, so dass ein Hamilton-Kreis für jede Variable einen Durchlauf auswählt.



Als nächstes konstruieren wir die Repräsentation der Klauseln. Für jede Klausel in F wird ein Teilgraph wie unten hinzugefügt.

Dieser hat 3 Ein- und Ausgänge, für jedes Literal in der Klausel jeweils einen. Wesentliche Eigenschaft ist, dass ein Hamilton-Kreis, der den Teilgraphen über einen Eingang erreicht, ihn nur über den gegenüberliegenden Ausgang verlassen kann, das sonst Knoten nicht mehr erreichbar sind. Dabei können einer oder zwei der anderen Ein- und Ausgänge mit abgedeckt werden.



Die Ein- und Ausgänge werden nun mit den Variablengraphen verbunden. Ist z.B das erste Literal x_i, dann wird wie gezeigt mit den schraffierten Knoten oben verbunden.

Eine Bewertung, die F erfüllt, entspricht dann genau einem Hamilton-Kreis, der die Variablen-Teilgraphen entsprechend der Werte durchläuft und dabei die Kausel-Teilgraphen genau an den mit 1 bewerteten Literalen betritt und verlässt.

Das Problem des Handlungsreisenden

Problem TSP

Instanz: $n \times n$-Matrix $D = \big(d(i,j)\big)$ mit $d(i,j) \in \mathbb{N}$,
Zahl $k \in \mathbb{N}$

Frage: Gibt es eine Permutation $p_1, \ldots, p_n$
mit $\sum_{i=1}^{n-1} d(p_i, p_{i+1}) + d(p_n, p_1) \leq k$.

Satz

TSP *ist NP-vollständig*

Beweis: durch Reduktion von Hamilton-Kreis auf TSP.

Zur Motivation des Probems TSP stellen wir uns einen Handlungsreisenden vor, der n Ständte bereisen soll. In der Matrix sind im Eintrag $d(i,j)$ die Kosten für die Reise von Stadt i nach Stadt j beschrieben.

Das Problem ist nun, eine Reihenfolge der Städte zu finden, die die gesamten Reisekosten minimiert, bzw. als Entscheidungsproblem: gibt es eine Reihenfolge, in der ich alle Städte mit Budget k bereisen kann.

Bei der Reduktion von Hamilton-Kreis auf TSP bildet man die n Knoten des Graphen auf n Städte ab. Wenn der Graph eine Kante zwischen i und j hat, macht man die Kosten von i nach j sehr klein, also $d(i,j) = 1$. Für Knoten i und j, zwischen denen keine Kante ist, macht man die Kosten sehr teuer, etwa $d(i,j) = n + 1$.

Eine Rundreise, die mit Kosten n auskommt, muss also nur Kanten im Graphen entlang laufen, und entspricht somit genau einem Hamilton-Kreis, und umgekehrt.

Das Problem SUBSET SUM

Problem SUBSET SUM

Instanz:	Zahlen $a_1, a_2, \ldots, a_n \in \mathbb{N}$, $t \in \mathbb{N}$
Frage:	Gibt es eine Teilmenge $I \subseteq \{1, \ldots, n\}$ mit $\sum_{i \in I} a_i = t$?

Theorem

SUBSET SUM *ist NP-vollständig.*

Das Problem ist offensichtich in NP.

Wir zeigen: 3-SAT $\leq_P$ SUBSET SUM.

Zur Motivation des Problems SUBSET SUM betrachte folgendes Problem:

Man hat ein System mit zwei Verarbeitungseinheiten, auf denen n Aufträge abgearbeitet werden sollen. Die Zeit zur Verarbeitung des i-ten Jobs ist a_i. Sei $t \geq (1/2) \sum_{i=1}^{n} a_i$.

Dann kann man die Jobs so auf die zwei Einheiten verteilen, dass die Gesamtverarbeitungszeit t ist, wenn das Problem SUBSET SUM mit Eingabedaten $a_1, \ldots, a_n$ und t eine Lösung hat. Dabei werden die Jobs i, für die a_i in der ausgewählten Menge ist, auf der einen und die übrigen auf der anderen Einheit verarbeitet.

Reduktion von 3-SAT auf SUBSET SUM

Sei $F = C_1 \wedge \ldots \wedge C_m$ eine Formel in 3-KNF, in den Variablen $x_1, \ldots, x_n$

Für $1 \leq i \leq n$:

$$a_i := 10^{m-1+i} + \sum_{x_i \in C_j} 10^{j-1}$$
$$b_i := 10^{m-1+i} + \sum_{\bar{x}_i \in C_j} 10^{j-1}$$

Für $1 \leq j \leq m$:

$$c_j = d_j = 10^{j-1}$$

$$t := \sum_{i=1}^{n} 10^{m-1+i} + \sum_{j=1}^{m} 3 \cdot 10^{j-1}$$

Es gilt: Es gibt eine Teilmenge der Zahlen $a_1, b_1, \ldots, a_n, b_n$, $c_1, d_1, \ldots, c_m, d_m$ mit Summe t gdw. F erfüllbar ist.

Die Idee der Konstruktion ist: die Stellen 10^0 bis 10^{m-1} entsprechen den Klauseln $C_1, \ldots, C_m$, und die Stellen 10^m bis 10^{m+n-1} den Variablen $x_1, \ldots, x_n$. a_i hat eine 1 in der Stelle $m-1+i$, sowie in jeder Stelle $j < m$ wo x_i in der Klausel C_{j+1} vorkommt. Ebenso hat b_i eine 1 in der Stelle $m-1+i$, sowie in jeder Stelle $j < m$ wo $\bar{x}_i$ in der Klausel C_{j+1} vorkommt.
In jeder der vorderen n Stellen stehen insgesamt bei allen Zahlen a_i, b_i und c_j, d_j nur 2 Einser, und in den hinteren m Stellen höchstens 5 Einser, da F eine 3-KNF-Formel ist. Daher kann stellenweise ohne Übertrag addiert werden.

Ist nun α eine Bewertung mit $\alpha \models F$, dann wählt man a_i aus wenn $\alpha(x_i) = 1$ ist und b_i wenn $\alpha(x_i) = 0$ ist. Damit hat die Summe in jeder der vorderen n Stellen den Wert 1. Da α jede Klausel erfüllt, hat in der Summe jede der hinteren m Stellen einen Wert zwischen 1 und 3. Daher wählt man noch so viele der Zahlen c_j, d_j aus, dass der Wert überall 3 wird, und kommt so auf die Summe t.

Hat man umgekehrt eine Teilmenge mit Summe t, so muss für jedes $1 \leq i \leq n$ genau eine der Zahlen a_i, b_i dabei sein, um in der Stelle $m-1+i$ auf den Wert 1 zu kommen. Daraus erhält man eine Bewertung, die diejenigen x_i auf 1 setzt, für die a_i dabei ist, und die anderen auf 0.
In den Stellen $0 \leq j \leq m-1$ kann durch die Zahlen c_j, d_j höchstens ein Wert 2 zusammenkommen, also muss eine der a_i, b_i einen Beitrag leisten. Dann wird aber die Klausel C_{j+1} durch $\alpha(x_i)$ erfüllt, weil sie entsprechend x_i oder $\bar{x}_i$ enthält.