# Komplexitätstheorie

## WiSe 2011/12 in Trier

Henning Fernau
Universität Trier
fernau@uni-trier.de

## Komplexitätstheorie Gesamtübersicht

- Organisatorisches / Einführung
  Motivation / Erinnerung / Fragestellungen

- Diskussion verschiedener Komplexitätsklassen:
  Zeitkomplexität
  Platzkomplexität

- zugehörige Reduktionsbegriffe

- vollständige Probleme

- Anpassung von Klassenbegriffen und Reduktionen

## Umfangreiche Liste NP-vollständiger Probleme

- Garey, M.R., Johnson, D.S., Computers and Intractability:
  A Guide to the Theory of **NP**-Completeness, Freeman, San Francisco 1979
  sowie als Fortsetzung davon:

- Johnson, D.S., The **NP**-completeness column: an ongoing guide,
  seit 1981 in der Zeitschrift "Journal of Algorithms",
  später in "ACM Transactions on Algorithms"

Noch einmal als Merksatz:

Solange es nicht gelungen ist, $\mathbf{P} = \mathbf{NP}$ zu beweisen, ist für keines der **NP**-vollständigen Probleme ein praktisch verwendbarer Algorithmus bekannt!

Im Folgenden **NP**-vollständige Probleme aus verschiedenen Problembereichen:

Logik, Graphentheorie, Mengentheorie und Zahlentheorie

Logik: $\rightsquigarrow$ Letzte Vorlesung
Graphentheorie / Mengentheorie: heute

## Probleme aus Graphentheorie:

**Satz 1** $\mathbf{NP}$*-vollständig (bzgl.* $\leq_{\log}$*) sind:*

- SIMPLE MAX CUT
- VERTEX COVER
- CLIQUE
- INDEPENDENT SET
- 3-FÄRBBARKEIT
- GERICHTETER HAMILTON-KREIS
- UNGERICHTETER HAMILTON-KREIS
- TRAVELING SALESMAN
- SUBGRAPH ISOMORPHISM

SIMPLE MAX CUT:

Gegeben: ein ungerichteter Graph $G = (V, E)$ und eine natürliche Zahl $k$.

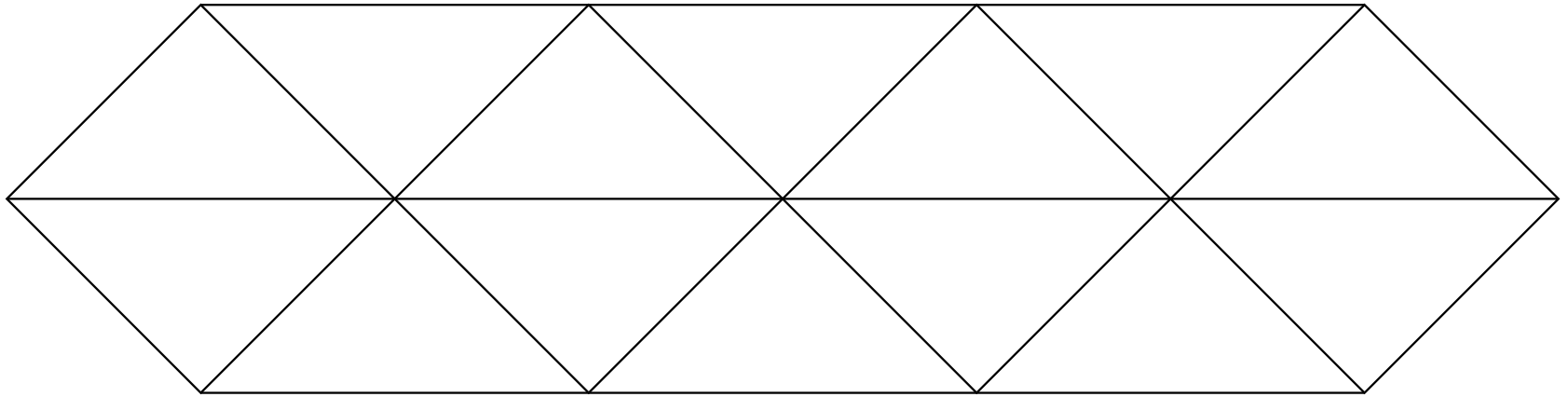
Frage: Kann man die Knotenmenge $V$ so in zwei disjunkte Teilmengen $V_a, V_b$ aufteilen, dass die Zahl der Kanten, die ein Ende in $V_a$ und ein Ende in $V_b$ haben, mindestens $k$ ist?

VERTEX COVER:

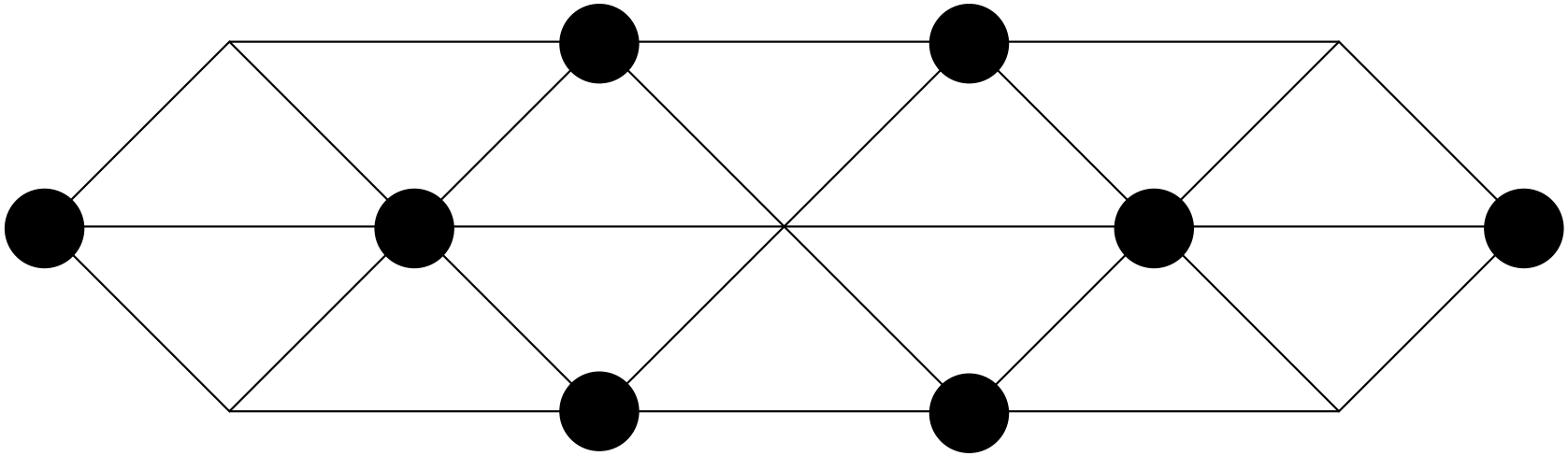
Gegeben: ein ungerichteter Graph $G = (V, E)$ und eine natürliche Zahl $k$.

Frage: Gibt es eine Menge $V' \subseteq V$ aus höchstens $k$ Knoten, die eine Knotenüberdeckung bildet?

Beispiel für die Begriffe: Betrachte folgenden Graphen:

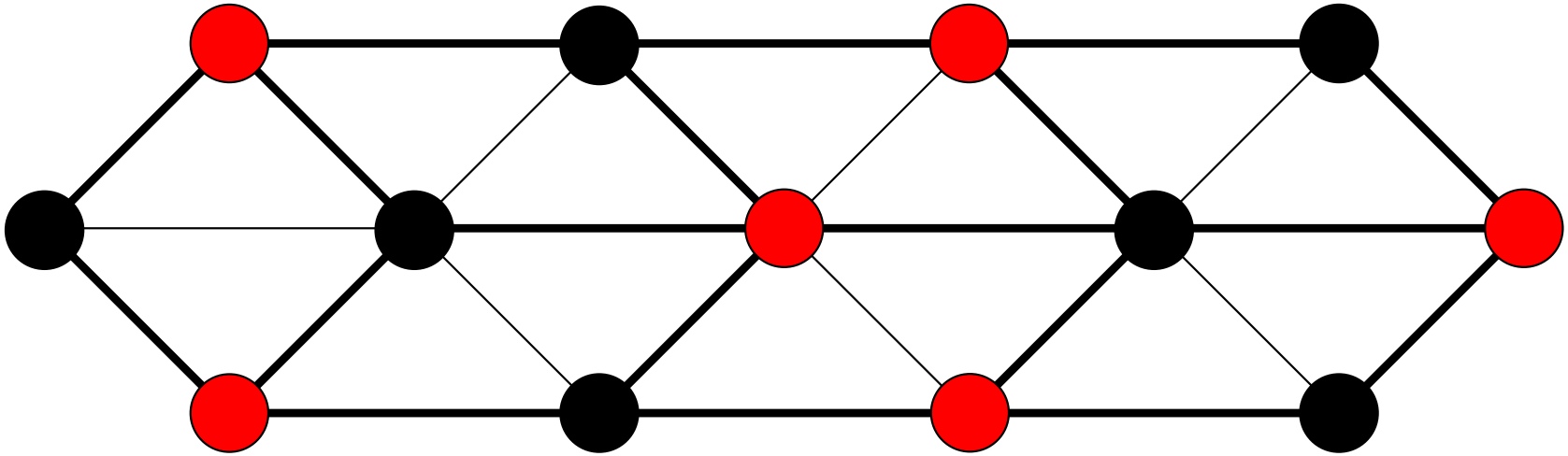


Beispiel: Lösung zu VERTEX COVER mit $k \geq 8$



Optimalitätsbeweis: Betrachte drei knotendisjunkte Dreiecke sowie zwei weitere disjunkte Kanten am Rand.

Beispiel: Lösung zu SIMPLE MAX CUT mit $k = 19$



Graph enthält 7 kantendisjunkte Dreiecke

$\rightsquigarrow$ höchstens zwei Kanten jedes dieser Dreiecke kommt in den Schnitt

$\rightsquigarrow$ $k = 19$ ist maximal, da 26 Kanten insgesamt

Sei $G = (V, E)$ ungerichteter Graph.
*unabhängige Knotenmenge*: Menge $V_1$ von Knoten, von denen keine zwei durch eine Kante verbunden sind.
*Clique* in $G$: Menge $V_2$ von Knoten, von denen je zwei durch eine Kante verbunden sind
(vgl. frühere Vorlesung: `CLIQUE(k)` und `INDEPENDENT SET(k)`)

- `CLIQUE`: Gegeben: ungerichteter Graph $G = (V, E)$ und $k > 0$.
  Frage: Gibt es in $G$ eine Clique mit mindestens $k$ Elementen?
- `INDEPENDENT SET`: Gegeben: ungerichteter Graph $G = (V, E)$ und $k > 0$.
  Frage: Gibt es in $G$ eine unabhängige Knotenmenge mit mindestens $k$ Elementen?

3-FÄRBBARKEIT: Gegeben: ungerichteter Graph $G = (V, E)$.

Frage: Gibt es eine Färbung der Knotenmenge $V$ mit drei Farben, sodass keine benachbarten Knoten die gleiche Farbe erhalten?

*Färbung*: Abbildung $Fb : V \to \{0, 1, 2\}$

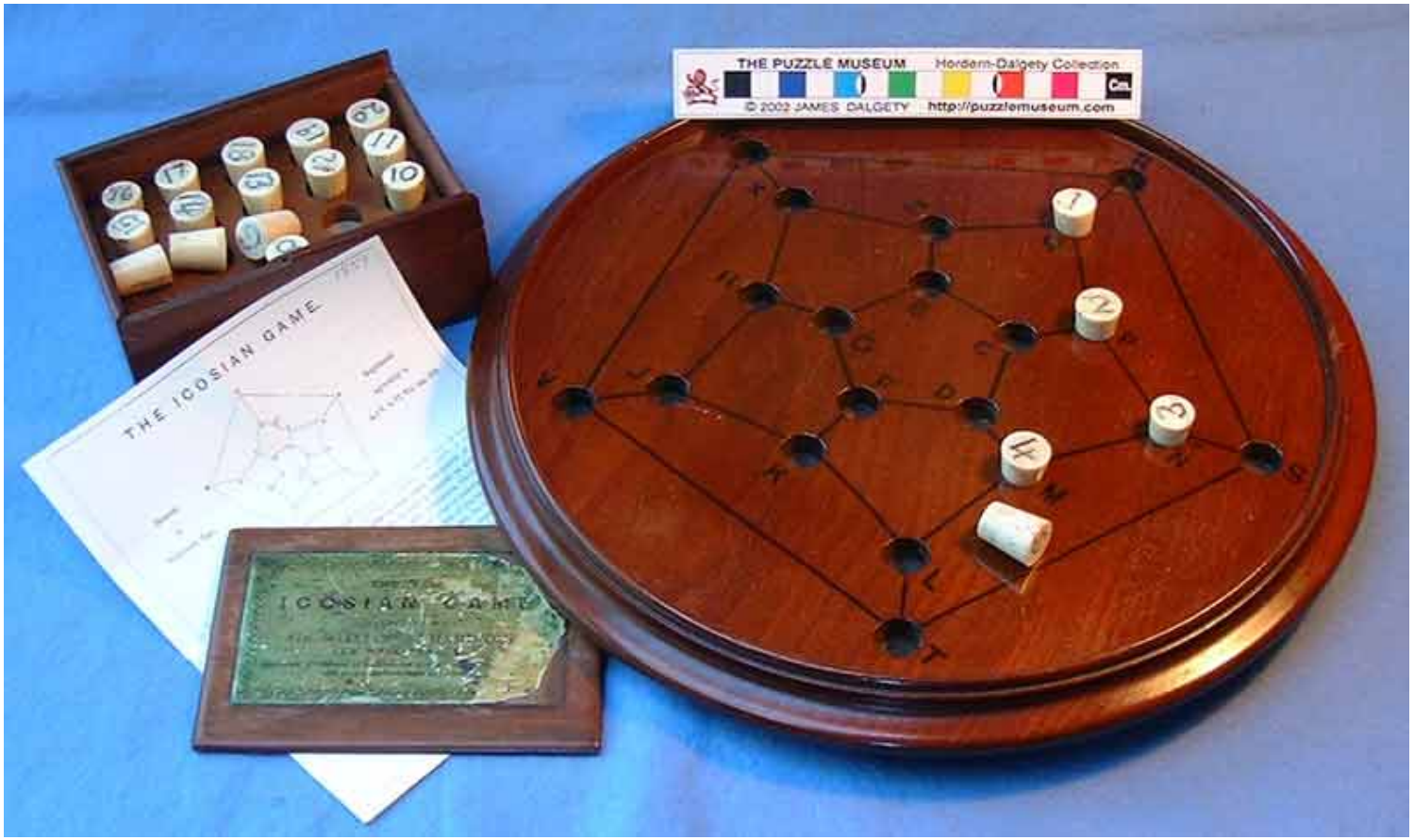
Verallgemeinerung auf $k$ Farben:
$k$-FÄRBBARKEIT **NP**-vollständig für $k \geq 3$
$k = 2$: schnell deterministisch lösbar ist, in $co - \mathbf{NL}$.
Umformulierung von 2-FÄRBBARKEIT:
Gibt es *keinen* Kreis mit ungerader Länge im Graphen?
Lewis und Papadimitriou (Symmetric Space-Bounded Computation, Theor. Comput. Sci. 19, 1982) sowie Reingold (J. ACM 55, 2008) haben gezeigt, das 2-FÄRBBARKEIT L-vollständig ist.

*Kreise* im Graphen:
—Pfad, der in dem Knoten endet, in dem er auch beginnt
—Knoten $x_1, \ldots x_m$ mit $x_m = x_1$ und $(x_i, x_{i+1}) \in E$ für $1 \leq i < m$ (bzw. $\{x_i, x_{i+1}\} \in E$ bei ungerichteten Graphen).

- GERICHTETER HAMILTON-KREIS:
  Gegeben: gerichteter Graph $G = (V, E)$.
  Frage: Gibt es in $G$ Kreis, der alle Knoten genau einmal berührt, d.h. gibt es Permutation $(x_1, \ldots, x_n)$ von $V$, für die $x_1, \ldots, x_n, x_1$ ein Kreis ist?

- UNGERICHTETER HAMILTON-KREIS:
  Gegeben:ungerichteter Graph $G = (V, E)$.
  Frage: Gibt es in $G$ Kreis, der alle Knoten genau einmal berührt?

**Einschub:** Mathematische Spiele; Gelderwerb für Hamilton



- TRAVELING SALESMAN:
  Gegeben: ungerichteter Graph $G = (V, E)$, Funktion $f : E \to \mathbb{N}$, Zahl $k \in \mathbb{N}$.
  Frage: Gibt es in $G$ Kreis $x_1, \ldots x_m, x_1$, der alle Knoten mindestens einmal berührt, sodass die Summe aller $f(e)$ für die Kanten $e$ im Kreis höchstens $k$ ergibt, d.h. folgender Ungleichung genügt?

$$\sum_{1 \leq i \leq m} f(x_i, x_{i+1}) \leq k$$

Anmerkungen:
—oft 'Kreis' so definiert: kein Knoten mehrfach durchlaufen (einfacher Kreis)!
—im Pfad nur Anfangs- und Endknoten identisch
—unwichtig bei HAMILTON-Problemen
—wichtig bei TRAVELING SALESMAN (in Literatur ohnehin unterschiedlich definiert!)

$G = (V, E)$ ist ein *Teilgraph* von $G' = (V', E')$, wenn $V \subseteq V'$ und $E \subseteq E'$, $E$ enthält nur Endpunkte aus $V$.

$G$ ist *isomorph* zu $G'' = (V'', E'')$, wenn Bijektion $f\colon V \to V''$ mit $(x, y) \in E \iff (f(x), f(y)) \in E''$ existiert.

- SUBGRAPH ISOMORPHISM:
  Gegeben: zwei gerichtete (oder ungerichtete) Graphen $G', G''$
  Frage: Gibt es einen Teilgraphen $G$ von $G'$, der zu $G''$ isomorph ist?

Hier nur Teil der Vollständigkeitsbeweise, Rest: Übungen

## Zu den Übungen 1

Vollständigkeitsbeweise für ein Problem $\mathcal{P}$ erfordern immer zwei Beweisschritte:

- Enthaltensein in NP
- Härte für NP

Für den ersten Schritt: entweder Angabe eines nichtdeterministischen Algorithmus oder Angabe einer geeigneten Reduktion, d.h., zeige $\mathcal{P} \leq_{\log} \mathcal{Q}$ für ein geeignet gewähltes Problem $\mathcal{Q}$ aus NP.
Für den zweiten Schritt: Wähle geeignetes NP-hartes Problem $\mathcal{H}$ und zeige $\mathcal{H} \leq_{\log} \mathcal{P}$.

## Zu den Übungen 2

Die Korrektheit einer Reduktion $\mathrm{r}$ für $\mathcal{R} \leq_{\log} \mathcal{S}$ erfordert vier Überlegungen:

1. Jeder Instanz $\mathrm{I}$ von $\mathcal{R}$ wird eine Instanz $\mathrm{r}(\mathrm{I})$ von $\mathcal{S}$ zugeordnet.

2. $\mathrm{r}$ kann mit einer logspace-Maschine berechnet werden.

3. Ist $\mathrm{I}$ eine JA-Instanz von $\mathcal{R}$, so ist $\mathrm{r}(\mathrm{I})$ eine JA-Instanz von $\mathcal{S}$.

4. Ist $\mathrm{r}(\mathrm{I})$ eine JA-Instanz von $\mathcal{S}$, so ist $\mathrm{I}$ eine JA-Instanz von $\mathcal{R}$.

Schwierigkeiten bei der erdachten Konstruktion zeigen sich zumeist im vierten Schritt.

## $3\text{-SAT} \leq_{log}$ GERICHTETER HAMILTON-KREIS (GHK)

Sei $w = (C_1 \wedge \cdots \wedge C_m)$ ein 3-CNF-Ausdruck mit Variablen $x_1, \ldots, x_n$.
Konstruiere Graph $G = (V, E)$ mit $w \in 3-\text{SAT} \iff G \in \text{GHK}$.
Insbesondere: Variable sind Knoten von $G$.

Erwünschte Eigenschaften:

—Es gibt von $x_i$ nach $x_{(i \bmod n)+1}$ stets zwei mögliche Pfade, einer für "$x_i$ ist wahr" und einer für "$x_i$ ist falsch".
—Für jede Klausel $C_j$ enthält $G$ Teilgraph ("Gadget") $H_j$, sodass bei "richtiger" Wahl der Pfade für $x_i$ (wahr/falsch gemäß der Interpretation $\phi$) alle Knoten in $H_j$ genau dann im Kreis $x_1 \to \cdots \to x_1$ durchlaufen werden können, wenn $C_j = (a_j \vee b_j \vee c_j)$ ein bzgl. $\phi$ wahres Literal enthält.

**Der Teilgraph $H_j$ für $C_j = (a_j \vee b_j \vee c_j)$**

Hinweis: Auf der nächsten Folie andere Variablenbezeichnungen wie $x_i(j)$.

Je zwei Knoten $A_j$, $A'_j$ je Literal $a_j$ etc.

Drei Arten gerichteter Kanten:

(1) $A_j \to B_j \to C_j \to A_j$

(2) $C'_j \to B'_j \to A'_j \to C'_j$

(3) $A_j \to A'_j$, $B_j \to B'_j$, $C_j \to C'_j$.

Außerdem Anbindungen von / nach "außen" über "$a_j$ wahr"-Pfade etc.

Erwünschte Eigenschaften:

—$H_j$ muss bei jeder nichtleeren Teilmenge $L$ von $\{a_j, b_j, c_j\}$ ganz durchlaufen werden, wobei: $\phi$ weist (genau) jedem $\ell \in L$ "wahr" zu.

—Die "Schnittstelle" zu den anderen Teilgraphen muss stimmen:
Nach Durchlauf durch $H_j$ müssen die gleichen $x_i$-Belegungspfade aus $H_j$ heraus- ausführen, wie sie hineingelaufen sind.

## Umsetzung der Eigenschaften von $H_j$:

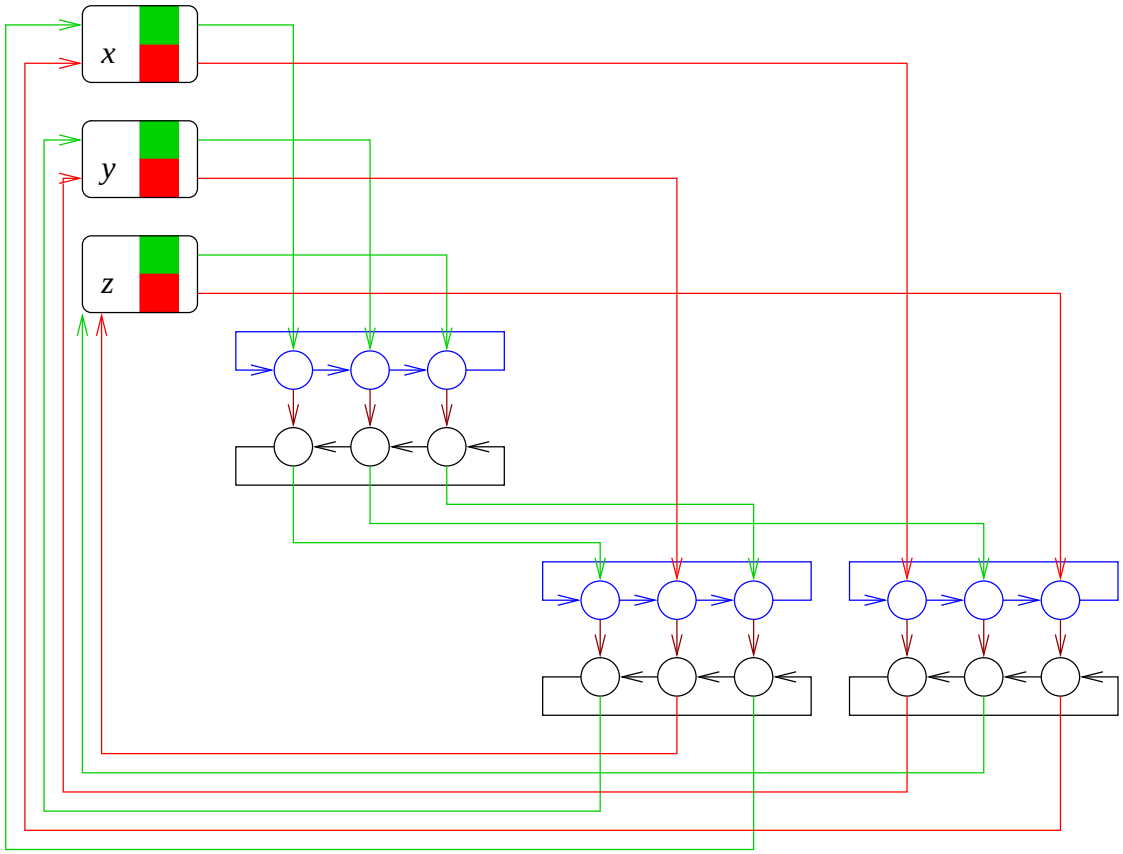
| Eingang | Durchlauf | Ausgang |
|---|---|---|
| $a_j$ wahr | $A_j \to B_j \to C_j \to C_j' \to B_j' \to A_j'$ | $A_j'$ wahr |
| $b_j$ wahr | $B_j \to C_j \to A_j \to A_j' \to C_j' \to B_j'$ | $B_j'$ wahr |
| $c_j$ wahr | $C_j \to A_j \to B_j \to B_j' \to A_j' \to C_j'$ | $C_j'$ wahr |
| $a_j, b_j$ wahr | $A_j \to A_j'; B_j \to C_j \to C_j' \to B_j'$ | $A_j', B_j'$ wahr |
| $a_j, c_j$ wahr | $C_j \to C_j'; A_j \to B_j \to B_j' \to A_j'$ | $A_j', C_j'$ wahr |
| $b_j, c_j$ wahr | $B_j \to B_j'; C_j \to A_j \to A_j' \to C_j'$ | $B_j', C_j'$ wahr |
| $a_j, b_j, c_j$ wahr | $A_j \to A_j'; B_j \to B_j'; C_j \to C_j'$ | $A_j', B_j', C_j'$ wahr |

Man mache sich klar: Es gibt keine anderen Möglichkeiten, durchlaufende “wahr-Pfade” zu konstruieren.

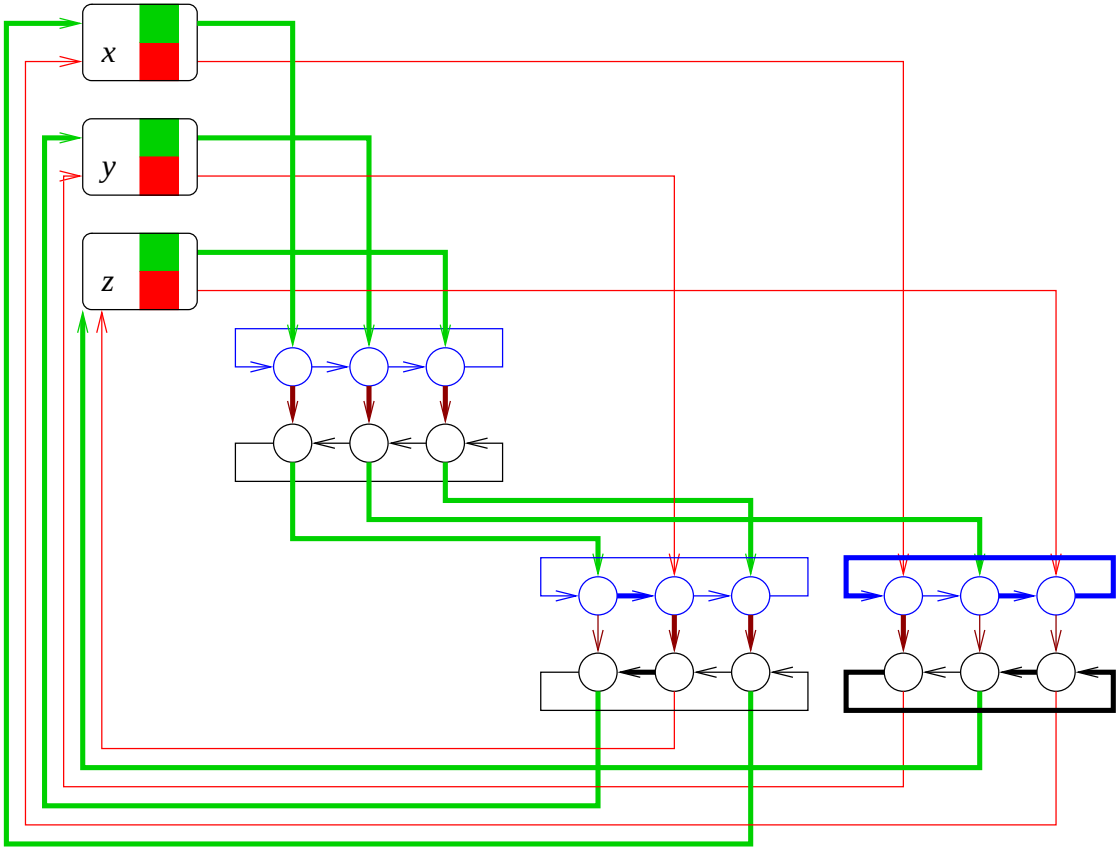
## Gesamtkonstruktion

Knotenmenge enthält: (1) $\{x_1, \ldots, x_n\}$
(2) $\{x_i(j), x_i'(j) \mid 1 \leq i \leq n, 1 \leq j \leq m, x_i \in C_j\}$
(3) $\{\overline{x_i}(j), \overline{x_i}'(j) \mid 1 \leq i \leq n, 1 \leq j \leq m, \overline{x_i} \in C_j\}$
Kantenmenge enthält:
(1) $H_j$-Gadget-Kanten für $C_j$ (das betrifft Verbindungen zwischen $x_i(j)$, $x_i'(j)$, $\overline{x_i}(j)$, $\overline{x_i}'(j)$)
(2) $(x_i, x_i(j))$, falls $C_j$ die erste Klausel in $w$ ist, die $x_i$ als Literal enthält (also: $x_i \in C_j$),
$(x_i, \overline{x_i}(j))$, falls $C_j$ die erste Klausel in $w$ ist mit $\overline{x_i} \in C_j$
(3) $(x_i'(j), x_i(k))$, falls $C_k$ die erste Klausel in $w$ nach $C_j$ ist mit $x_i \in C_k$,
$(\overline{x_i}'(j), \overline{x_i}(k))$, falls ...
(4) $(x_i'(j), x_{(i \bmod n)+1})$, falls $C_j$ die letzte Klausel in $w$ ist mit $x_i \in C_j$,
$(\overline{x_i}'(j), x_{(i \bmod n)+1})$, falls ...

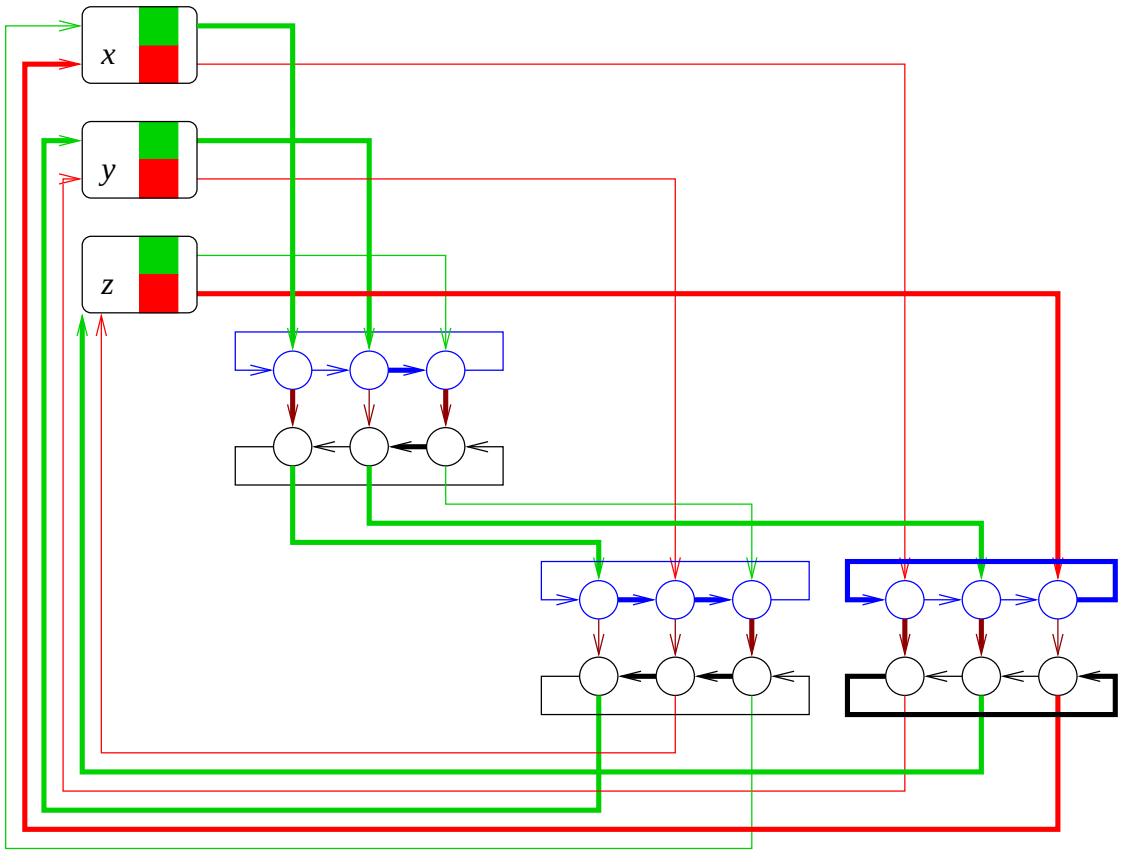
## Ein Beispiel für die Konstruktion $w = ((x \vee y \vee z) \wedge (x \vee \bar{y} \vee z) \wedge (\bar{x} \vee y \vee \bar{z}))$



## Eine Lösung für das Beispiel $\phi(x) = \phi(y) = \phi(z) = 1$.



## Eine weitere Lösung für das Beispiel $\phi(x) = \phi(y) = 1;\ \phi(z) = 0.$



## Ein kleines Eingeständnis

Wie oft, wird bei der Konstruktion ein wenig “geschludert”:
Implizit wird bei der Kantenangabe in (2) und (4) davon ausgegangen, dass die erwähnten “ersten” bzw. “letzten” Klauseln mit $x_i$ überhaupt existieren.

Wie kann man die Konstruktion “retten” ?
**Möglichkeit 1**: Genauere Beschreibung der Kantenmenge, die “wirklich” gemeint ist.
**Möglichkeit 2**: Was bedeutet es auf der logischen (SAT) Ebene, wenn es gar keine Klauseln mit $x_i$ gibt ?
Gibt es vielleicht eine Normalform für 3-SAT, die diesen Fall ausschließt ?
Gibt es ein Polynomzeitverfahren, das diese Normalform herstellt ?
In welchem Sinne könnte man dies als Reduktion auffassen ?

Welche weiteren Mängel an der beschriebenen Reduktion (und der Beweisskizze) sind Ihnen aufgefallen ?

## $\mathrm{GHK} \leq_{\log}$ UNGERICHTETER HAMILTON-KREIS (UHK)

Sei $G = (V, E)$ gerichteter Graph.

Konstruiere ungerichteten Graphen $G' = (V', E')$ wie folgt:

—$V' = \{x_a, x_m, x_e \mid x \in V\}$,

—$E' = \{\{x_a, x_m\}, \{x_m, x_e\} \mid x \in V\} \cup \{\{x_e, y_a\} \mid (x, y) \in E\}$.

Gilt $G \in \mathrm{GHK}$, so gibt es Kreis $x_1, \ldots, x_n, x_1$ in $G$ mit $V = \{x_1, \ldots, x_n\}$.
Dann ist $x_{1a}, x_{1m}, x_{1e}, x_{2a}, \ldots, x_{na}, x_{nm}, x_{ne}, x_{1a}$ Kreis in $V'$, der alle Knoten von $V'$ durchläuft, d.h., $G' \in \mathrm{UHK}$.

Für jeden Kreis in $G'$, der alle Knoten durchläuft, gilt:
$x_m$ wird entweder über $x_a, x_m, x_e$ oder über $x_e, x_m, x_a$ erreicht.
Nach Konstruktion gibt es von $x_a$ nur Verbindungen zu $y_e$ und von $x_e$ nur Verbindungen zu $z_a$.
Also werden alle $x_m$ in gleicher Weise durchlaufen. O.E. hat der Kreis in $G'$ die folgende Bauart:
$z_{1a}, z_{1m}, z_{1e}, z_{2a}, \ldots, z_{ne}, z_{1a}$.
Dann ist $z_1, z_2, \ldots, z_n, z_1$ ein Hamilton-Kreis in $G$.

## Mengentheoretische Probleme

**Satz 2** **NP**-*vollständig (bzgl.* $\leq_{\log}$*) sind:*

- `TRIPARTITE MATCHING`
- `SET COVERING`
- `SET PACKING`
- `EXACT COVER BY 3 SETS`

- TRIPARTITE MATCHING:
  Gegeben: drei Mengen $\mathcal{M}$, $\mathcal{F}$, $\mathcal{H}$ mit gleicher Kardinalität

  $$\mathfrak{n} := \#\mathcal{M} = \#\mathcal{F} = \#\mathcal{H}$$

  und eine Menge $\mathcal{T}$ von Tripeln $\mathcal{T} \subseteq \mathcal{M} \times \mathcal{F} \times \mathcal{H}$.
  Frage: Gibt es eine Teilmenge $\mathcal{T}' \subseteq \mathcal{T}$ von $\mathfrak{n}$ Tripeln, so dass keine zwei Tripel eine gemeinsame Komponente haben?
  So eine Teilmenge heißt auch *(tripartites) / (dreidimensionales) Matching*.

- SET COVERING: Gegeben: Familie $F = \{S_1, ..., S_n\}$ von Mengen mit $S_i \subseteq U$ für ein endliches Universum $U$ sowie eine Zahl $k \in \mathbb{N}$.
  Frage: Gibt es $k$ Mengen $S_{i_1}, ..., S_{i_k}$ mit

$$\bigcup_{j=1}^{k} S_{i_j} = U$$

- SET PACKING: Gegeben: Familie $F = \{S_1, ..., S_n\}$ von Mengen mit $S_i \subseteq U$ für ein endliches Universum $U$ sowie eine Zahl $k \in \mathbb{N}$.
  Frage: Gibt es $k$ paarweise disjunkte Mengen $S_{i_1}, ..., S_{i_k}$ mit

$$\bigcup_{j=1}^{k} S_{i_j} = U$$

- EXACT COVER BY 3 SETS: Gegeben: Familie $\mathcal{F} = \{S_1, ..., S_n\}$ von Mengen mit $S_i \subseteq U$ mit $\#S_i = 3$ und $\#U = 3k$.
  Frage: Gibt es $k$ paaarweise disjunkte Mengen $S_{i_1}, ..., S_{i_k}$ mit

$$\bigcup_{j=1}^{k} S_{i_j} = U$$

Die Reduktionen

TRIPARTITE MATCHING $\leq_{\log}$ EXACT COVER BY 3 SETS

EXACT COVER BY 3 SETS $\leq_{\log}$ SET COVERING

EXACT COVER BY 3 SETS $\leq_{\log}$ SET PACKING

sind (fast) offensichtlich....

## Das Offensichtliche formaler...

TRIPARTITE MATCHING ist ein Spezialfall von EXACT COVER BY 3 SETS: Das Universum $U$ hat die besondere Eigenschaft, in drei gleich große Mengen $M, F, H$ zerlegt werden zu können, sodass jede Menge aus $\mathcal{F}$ jeweils ein Element dieser drei Mengen enthält.

EXACT COVER BY 3 SETS ist ein Spezialfall von SET COVERING mit $|U| = 3m$, $|S_i| = 3$ für alle $S_i \in \mathcal{F}$ und $k = m$.

Entsprechendes gilt für SET PACKING.

## 3-SAT $\leq$ TRIPARTITE MATCHING

Betrachte CNF-Formel $w = (C_1 \wedge C_2 \wedge \cdots \wedge C_m)$ mit Klauseln $C_j = (y_{j1} \vee y_{j2} \vee y_{j3})$ und Variablen $\{x_1, \ldots, x_n\}$.

Konstruiere 3 Mengen $W, X, Y$ gleicher Mächtigkeit $2nm$ und Tripelmenge $T \subseteq W \times X \times Y$, .
$W = \{u_{ij}^k \mid 1 \leq i \leq n, 1 \leq j \leq m, k = 0, 1\}$: Wahrheitswerte der Variablen $x_i$ bzgl. Klausel $C_j$ (sollte harmonisiert sein, s.u.).

$T$ beinhaltet drei Arten von Tripeln:

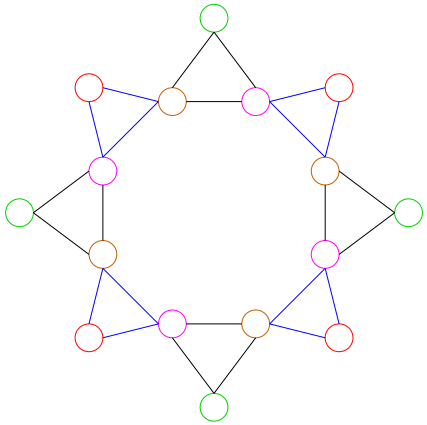
$H_i$ (Harmonisierung) Hat Variable $x_i$ in allen (simulierten) Klauseln denselben Wert?
$E_j$ (Erfüllungstest) Sind alle Klauseln $C_j$ erfüllt?
$G$ (Garbage Collection) Syntaktischer Müll, damit das Matching aufgeht.

$H_i$ Zu $X$ kommen Elemente $a_{ij}$ und zu $Y$ Elemente $b_{ij}$ hinzu ($1 \leq i \leq n, 1 \leq j \leq m$),
die in keinen Tripeln außer den Folgenden enthalten sind:

$H_i^1 = \{(u_{ij}^1, a_{ij}, b_{ij}) \mid 1 \leq j \leq m\}$.

$H_i^0 = \{(u_{ij}^0, a_{i,(j \bmod m)+1}, b_{ij}) \mid 1 \leq j \leq m\}$.

$H_i = H_i^1 \cup H_i^0$.



Jedes $H_i$ entspricht einem solchen “Stern” (hier $j = 4$).
Da $a_{ij}$, $b_{ij}$ (braun / magenta) sonst nicht vorkommen,
ist bei einem Matching $M'$ entweder $H_i^0$ oder $H_i^1$ enthalten
(blaue bzw. schwarze Dreiecke), aber niemals “gemischt”.
Die Wahrheitswertzuweisung kann z.B. mit

$x_i$ wahr gdw. $M' \cap H_i = H_i^1$

“abgelesen” werden (s. rot/grün-Färbung der Sternstrahlen)

$E_j$ zu $C_j = (y_{j1} \vee y_{j2} \vee y_{j3})$:
Zu $X$ kommen Elemente $s_{1j}$ und zu $Y$ Elemente $s_{2j}$ hinzu.
Ist $x_i$ das $k$-te Literal in $C_j$, so liegt $(u^0_{ij}, s_{1j}, s_{2j})$ in $T$.
Ist $\overline{x_i}$ das $k$-te Literal in $C_j$, so liegt $(u^1_{ij}, s_{1j}, s_{2j})$ in $T$.

Jedes Matching $M'$ enthält höchstens eines der drei Tripel $(??, s_{1j}, s_{2j})$.
Ist so ein Tripel enthalten, so entspricht dies der Erfüllung der Klausel $C_j$.
Genauer: Wurde durch $H_i$ $u^1_{ij}$ ausgewählt (also $x_i$ auf wahr gesetzt),
so ist $u^0_{ij}$ noch "frei" und kann zur "Erfüllung" von $C_j$ verwendet werden.

**Müll** $G$ Durch $\bigcup H_i$ und $\bigcup E_j$ sind von $W$ genau $mn + m$ Elemente abgedeckt.
$G$ muss die restlichen $(m-1)n$ Elemente abdecken:
$G = \{(u^k_{ij}, g_{1\ell}, g_{2\ell}) \mid 1 \leq i \leq n, 1 \leq j \leq m, 0 \leq k \leq 1, 1 \leq \ell \leq (m-1)n\}$.

Achtung: Konstruktion klappt nur, falls Variable nicht zweimal in irgendeiner Klausel vorkommt.