アナウンス（覚書）
- 講義のアンケート
- レポート返却
- 期末試験：5月31日チュートリアルアワー
  - 前回と同じ形式
  - 出題範囲は「計算量」のみ

I238 計算の理論

上原 隆平

2018年I-1期（4-5月）

I238 Computation Theory

by

Prof. Ryuhei Uehara

Term I-1, April-May, 2018

計算量の理論

- ゴール1:
  - “計算可能な*関数/問題/言語/集合*”

- ゴール2:
  - 「問題の困難さ」を示す方法を学ぶ
    - 計算可能な問題であっても、*手におえない*場合がある！
      - 計算に必要な資源（時間・領域）が多すぎるとき
    - 関連する専門用語;

      クラスNP, P≠NP予想, NP困難性, <u>還元</u>

Computational Complexity

- Goal 1:
  - *"Computable Function/Problem/Language/Set"*

- Goal 2:
  - How can you show *"Difficulty of Problem"*
    - There are *intractable* problems even if they are computable!
      - because they require too many resources (time/space)!
    - Technical terms;

      The class NP, P≠NP conjecture, NP-hardness, reduction

5.計算量の理論

- **計算量クラスの定義を概観すると…**

クラスPの定義
集合LがクラスPに入る ⇔
以下を満たす多項式時間計算可能述語Rが存在:
各 $x \in \Sigma^*$で$x \in L \Leftrightarrow R(x)$

クラスNPの定義
集合LがクラスNPに入る ⇔
以下を満たす多項式qと多項式時間計算可能述語Rが存在:
各 $x \in \Sigma^*$で$x \in L \Leftrightarrow \exists w \in \Sigma^*: |w| \leqq q(|x|)[R(x,w)]$

クラスcoNPの定義
集合LがクラスcoNPに入る ⇔
以下を満たす多項式qと多項式時間計算可能述語Rが存在:
各 $x \in \Sigma^*$で$x \in L \Leftrightarrow \forall w \in \Sigma^*: |w| \leqq q(|x|)[R(x,w)]$

5. Computational Complexity

- Observation of the classes

<u>Definition</u>: Class P

Set L is in the class P ⇔

There exists a poly-time computable predicate R such that

for each $x \in \Sigma^*$, $x \in L \Leftrightarrow R(x)$

<u>Definition</u>: Class NP

Set L is in the class NP ⇔

There exists a poly q and a poly-time computable predicate R such that

for each $x \in \Sigma^*$, $x \in L \Leftrightarrow \exists w \in \Sigma^* : |w| \leqq q(|x|)[R(x,w)]$

<u>Definition</u>: Class coNP

Set L is in the class coNP ⇔

There exists a poly q and a poly-time computable predicate R such that

for each $x \in \Sigma^*$, $x \in L \Leftrightarrow \forall w \in \Sigma^* : |w| \leqq q(|x|)[R(x,w)]$

6.多項式時間計算可能性の解析

6.2.完全性

(NP)完全性を示す二つの方法:

1. 定義に忠実に「すべての*L*'」に対して示す
   - クックの定理; 彼は1971年にSATでチューリングマシンのシミュレータを構築した！

基本的には…
1. 標準形で書かれたプログラムを
2. SATの命題論理式で模倣
   →非常に複雑＆面倒

例えば3SATは一様な構造を持っているので，扱いやすい．

2. すでに完全性が示されている問題をタネに使う
   - 3SAT$\leq_m^P$ DHAM, $\boxed{\text{3SAT} \leq_m^P \text{VC}}$, …
   - 千を超えるNP完全問題が3SATからの還元で示されている!
   - 例えば「一般のグラフ上でDHAMはNP完全」という結果から:
     - DHAMは平面グラフ上に限定してもNP完全
     - $\boxed{\text{DHAMは最大次数3に限定してもNP完全}}$
     - DHAMは二部グラフに限定してもNP完全…

最大次数5

6. Analysis on Polynomial-Time Computability

6.2. Completeness

There are two ways to prove (NP-)completeness:

1. show 'for all L' according to the definition
   - Cook's Theorem; he simulated Turing machine by SAT in 1971!

Basically…
1. For any program in standard form,
2. simulate it by SAT formulae
   →pretty complicated and tedious

Easy to handle since, e.g., 3SAT has a uniform structure.

2. use some known complete problem as a seed
   - 3SAT$\leq_m^P$ DHAM, $\boxed{\text{3SAT} \leq_m^P \text{VC},}$ …
   - Thousands of NP-complete problems are reduced from 3SAT!
   - E.g., from "DHAM is NP-complete for general graphs", we have
     - DHAM is NP-complete even for planar graphs
     - DHAM is NP-complete even for graphs with max degree=3
     - DHAM is NP-complete even for bipartite graphs…

max degree=5

6.多項式時間計算可能性の解析

6.2.完全性

定理 VCはNP完全問題

[証明] VC ∈ NPなので3SAT$\leq_m^P$ VCを示せばよい.

与えられた論理式 $F(x_1,x_2,\ldots,x_n)$ から以下の条件を満たすグラフと整数の組$\langle G,k \rangle$を多項式時間で構成する:

> $F()=1$ とする割当てが存在する
> ⇔G が大きさ k の頂点被覆をもつ

Gの構成方法 (F は n 変数・ m 項からなる):

1. Fの各変数 x_i に対して，頂点 x_i^+,x_i^- と辺 (x_i^+,x_i^-) を追加する
2. Fの各項 $C_j=(l_{i1} \vee l_{i2} \vee l_{i3})$ に対して，頂点 l_{i1}, l_{i2}, l_{i3} と3辺 (l_{i1},l_{i2}), (l_{i2},l_{i3}), (l_{i3},l_{i1}) を追加する
3. 各項 C_jに対して，リテラルl_{i1} が x_iなら辺 (l_{i1},x_i^+) を，$\neg x_i$ なら辺(l_{i1},x_i^-) を追加する
4. $k = n+2m$ とする

6. Analysis on Polynomial-Time Computability

6.2. Completeness

Theorem VC is NP-complete

[Proof] Since VC $\in$ NP, we show 3SAT $\leq_m^P$ VC.

For given formula $F(x_1,x_2,\ldots,x_n)$, we construct a pair $<G,k>$ of a graph and an integer in polynomial time such that:

$$\boxed{\begin{array}{l}\text{There is an assignment that makes } F()=1 \\ \Leftrightarrow G \text{ has a vertex cover of size } k\end{array}}$$

Construction of G (F has n variables and m clauses):

1. add vertices x_i^+, x_i^- and the edge (x_i^+, x_i^-) for each variable x_i in F
2. For each clause $C_j=(l_{i1} \vee l_{i2} \vee l_{i3})$ in F, add vertices l_{i1}, l_{i2}, l_{i3} and three edges (l_{i1},l_{i2}), (l_{i2},l_{i3}), (l_{i3},l_{i1})
3. add the edge (l_{i1},x_i^+) if the literal l_{i1} is x_i, or add (l_{i1},x_i^-) if it is $\neg x_i$ for each clause C_j
4. let $k = n+2m$

定理 VCはNP完全問題

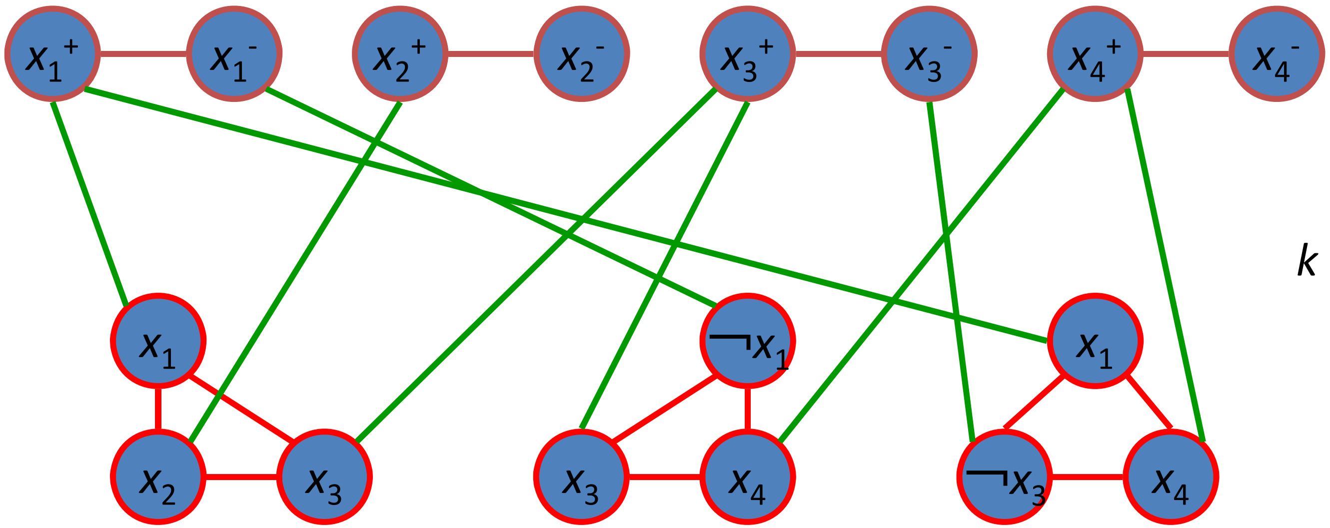
> $F()$=1 とする割当てが存在する
> ⇔G が大きさ k の頂点被覆をもつ

Gの構成方法 (F は n 変数・ m 項からなる):

1. Fの各変数 x_i に対して，頂点 x_i^+, x_i^- と辺 (x_i^+, x_i^-) を追加する
2. Fの各項 $C_j=(l_{i1} \vee l_{i2} \vee l_{i3})$ に対して，頂点 l_{i1}, l_{i2}, l_{i3} と3辺 (l_{i1}, l_{i2}), (l_{i2}, l_{i3}), (l_{i3}, l_{i1}) を追加する
3. 各項 C_jに対して，リテラルl_{i1} が x_iなら辺 (l_{i1}, x_i^+) を，$\neg x_i$ なら辺 (l_{i1}, x_i^-) を追加する
4. $k = n+2m$ とする

例: $F(x_1,x_2,x_3,x_4) = (x_1 \vee x_2 \vee x_3) \wedge (\neg x_1 \vee x_3 \vee x_4) \wedge (x_1 \vee \neg x_3 \vee x_4)$

G



k = 4+2×3=10

Theorem VC is NP-complete

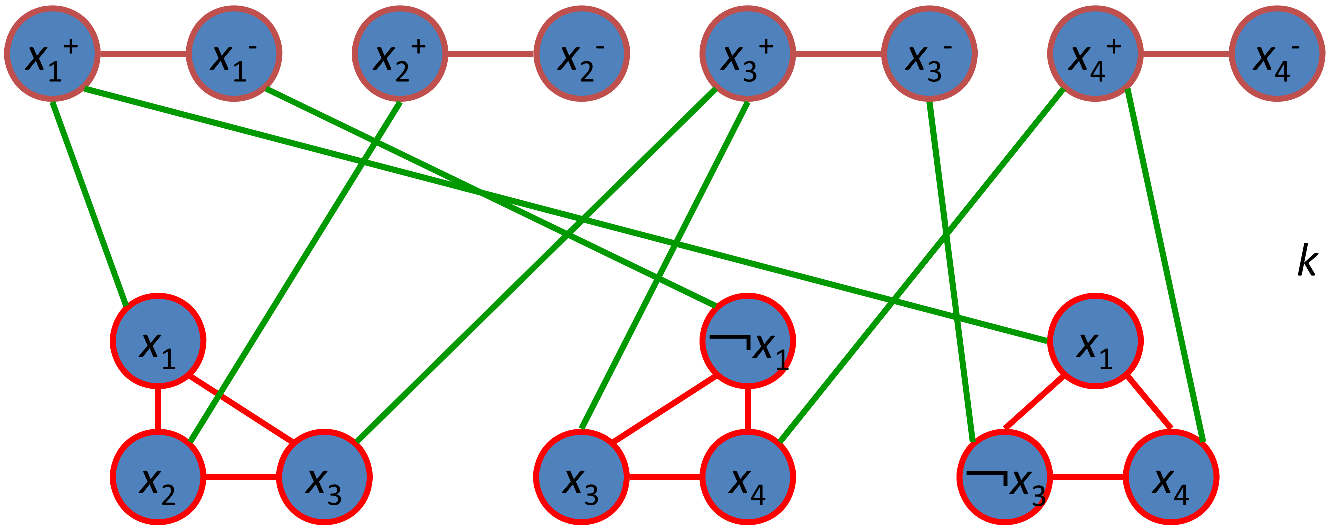
There is an assignment that makes $F()=1$
$\Leftrightarrow$ G has a vertex cover of size k

Construction of G (F has n variables and m clauses):

1. add vertices x_i^+, x_i^- and the edge (x_i^+, x_i^-) for each variable x_i in F
2. For each clause $C_j=(l_{i1} \vee l_{i2} \vee l_{i3})$ in F, add vertices l_{i1}, l_{i2}, l_{i3} and three edges (l_{i1}, l_{i2}), (l_{i2}, l_{i3}), (l_{i3}, l_{i1})
3. add the edge (l_{i1}, x_i^+) if the literal l_{i1} is x_i, or add (l_{i1}, x_i^-) if it is $\neg x_i$ for each clause C_j
4. let $k = n+2m$

Ex: $F(x_1,x_2,x_3,x_4) = (x_1 \vee x_2 \vee x_3) \wedge (\neg x_1 \vee x_3 \vee x_4) \wedge (x_1 \vee \neg x_3 \vee x_4)$

G



$k = 4+2 \times 3=10$

定理 VCはNP完全問題

FからGの構成は，明らかに多項式時間で可能である.
したがって，以下を証明すればよい：

> $F()=1$とする割当てがある
> ⇔G が大きさkの頂点被覆をもつ

観測：

Gの構成方法から，頂点被覆 S は以下の頂点を必ず含む

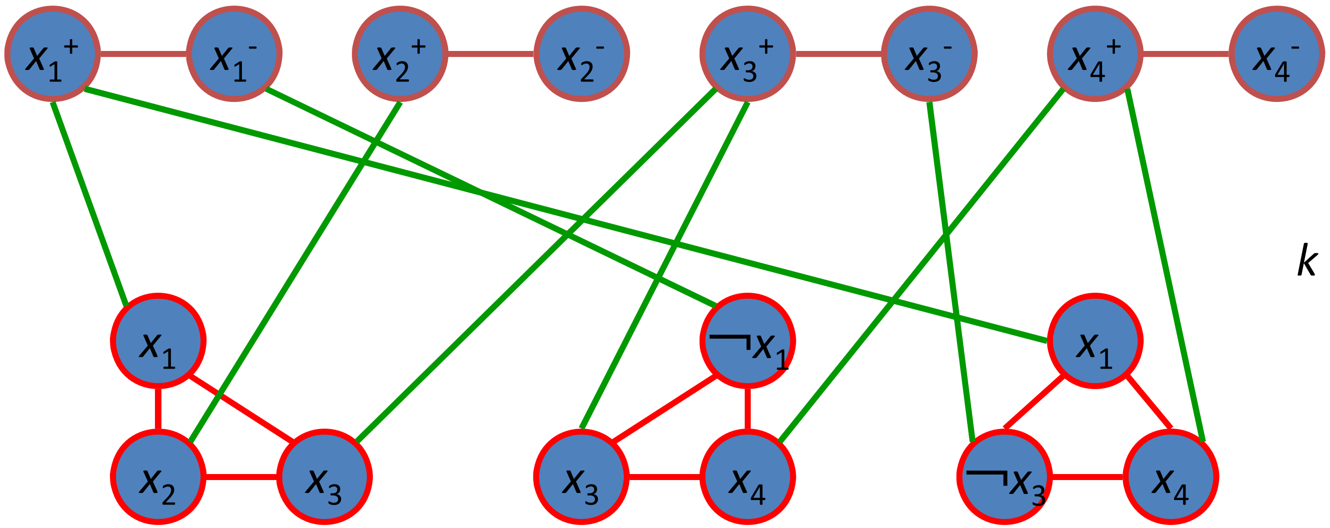
- x_i^+ or x_i^- から少なくとも一つ
- C_j の三つの頂点から少なくとも二つ

よって $|S| \geqq n+2m = k$

余分な頂点は一つもない！

例: $F(x_1,x_2,x_3,x_4) = (x_1 \vee x_2 \vee x_3) \wedge (\neg x_1 \vee x_3 \vee x_4) \wedge (x_1 \vee \neg x_3 \vee x_4)$

G



$k = 4+2\times 3=10$

Theorem VC is NP-complete

It is easy to see that the construction of G from F can be done in polynomial time of the size of F. Hence, we show that...

There is an assignment that makes $F()=1$ $\Leftrightarrow G$ has a vertex cover of size k

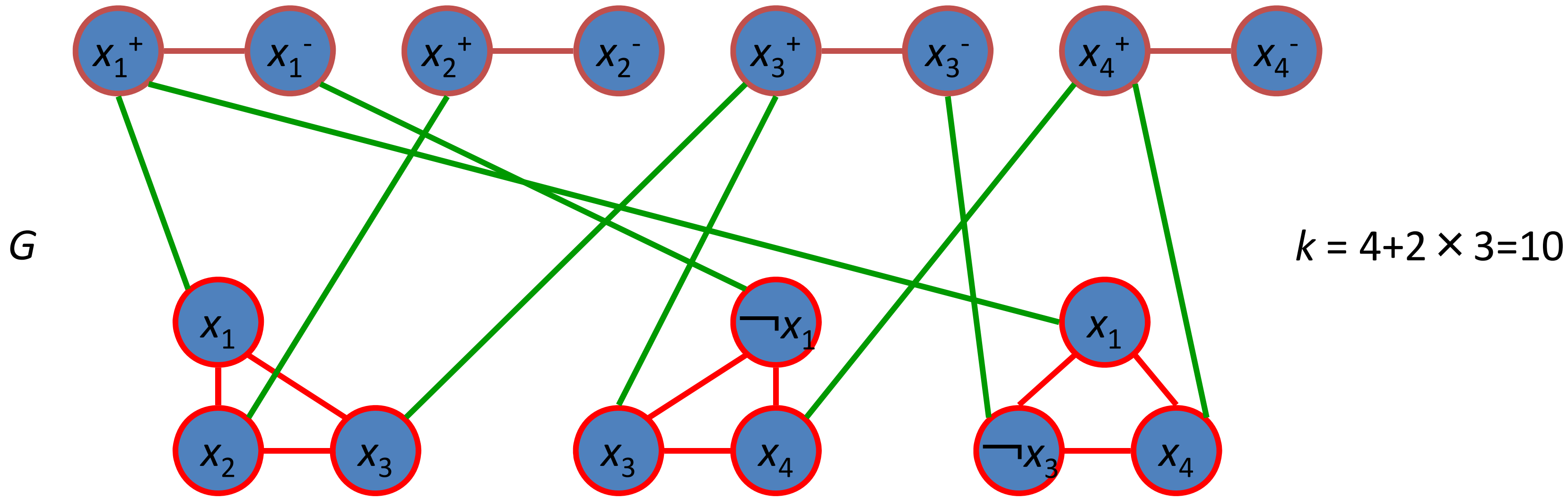
Observation:

From the construction of G, any vertex cover S should contain
- at least one of x_i^+ or x_i^-
- at least 2 of 3 vertices in C_j

Hence we have $|S| \geqq n+2m = k$.

We have no extra vertex!!

Ex: $F(x_1,x_2,x_3,x_4) = (x_1 \vee x_2 \vee x_3) \wedge (\neg x_1 \vee x_3 \vee x_4) \wedge (x_1 \vee \neg x_3 \vee x_4)$



定理 VCはNP完全問題

> $F()=1$とする割当てがある
> ⇒G が大きさkの頂点被覆をもつ

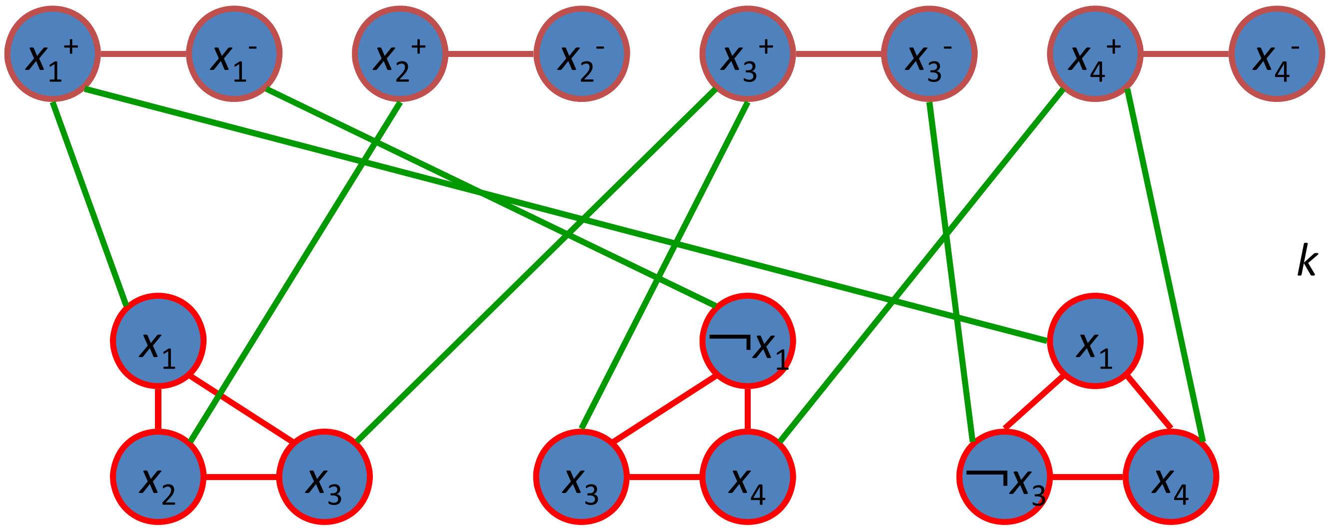
1.各 x_iに対して $\left\{ \begin{array}{l} x_i=1 \text{ なら } x_i^+ \\ x_i=0 \text{ なら } x_i^- \end{array} \right\}$ をSに

2. 各項 $C_j=(l_{i1},l_{i2},l_{i3})$ は充足されているので，少なくとも一つのリテラル l_{i1} に対して辺 (l_{i1},x_{i1}) は変数x_{i1}で被覆されている．そこで残りの二つのリテラル (l_{i2},l_{i3}) を S に

⇒ 観測 よりSは大きさkの頂点被覆．

例: $F(x_1,x_2,x_3,x_4) = (x_1 \vee x_2 \vee x_3) \wedge (\neg x_1 \vee x_3 \vee x_4) \wedge (x_1 \vee \neg x_3 \vee x_4)$

G



$k = 4+2 \times 3=10$

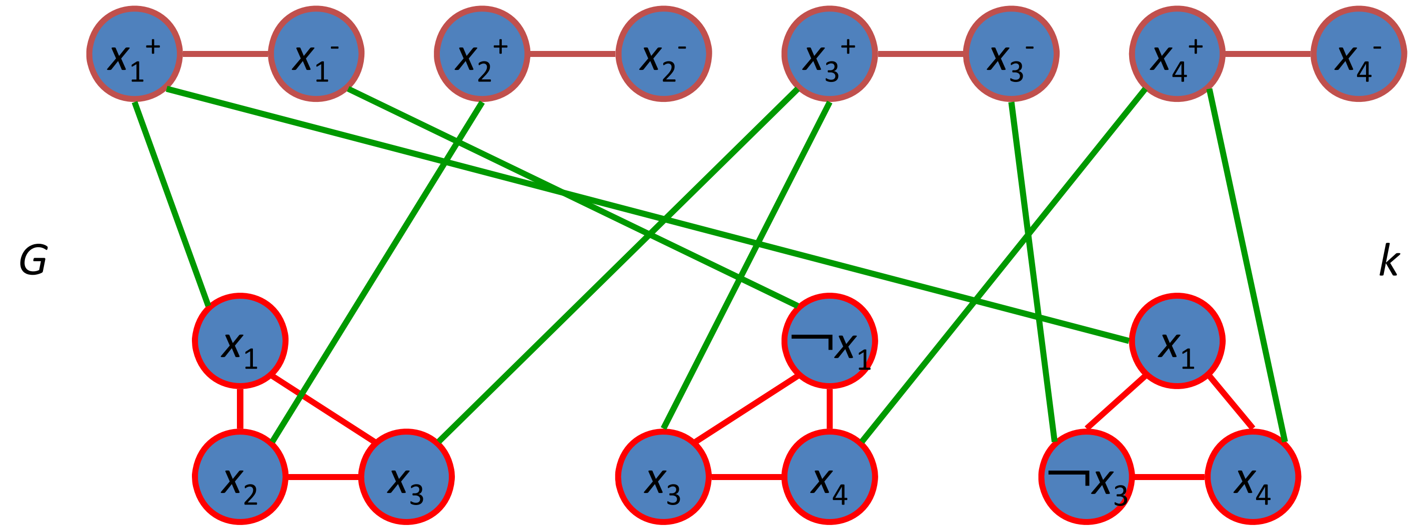
Theorem VC is NP-complete

There is an assignment that makes $F()=1$
$\Rightarrow G$ has a vertex cover of size k

1. Put $\left\{ \begin{array}{l} x_i^+ \text{ if } x_i=1 \\ x_i^- \text{ if } x_i=0 \end{array} \right\}$ into S for each x_i.
2. Since each clause $C_j=(l_{i1},l_{i2},l_{i3})$ is satisfied, at least one literal, say l_{i1}, the edge (l_{i1},x_{i1}) is covered by the variable x_{i1}. Therefore, put the remaining literals (l_{i2},l_{i3}) into S.

$\Rightarrow$ From the Observation S is a vertex cover of size k.

Ex: $F(x_1,x_2,x_3,x_4) = (x_1 \vee x_2 \vee x_3) \wedge (\neg x_1 \vee x_3 \vee x_4) \wedge (x_1 \vee \neg x_3 \vee x_4)$



$k = 4+2 \times 3=10$

定理 VCはNP完全問題

G が大きさkの頂点被覆をもつなら，$F()=1$とする割当てが存在する

1. 観測 より，被覆 S は各項から $2m$ 頂点含み，変数から n 頂点含む．

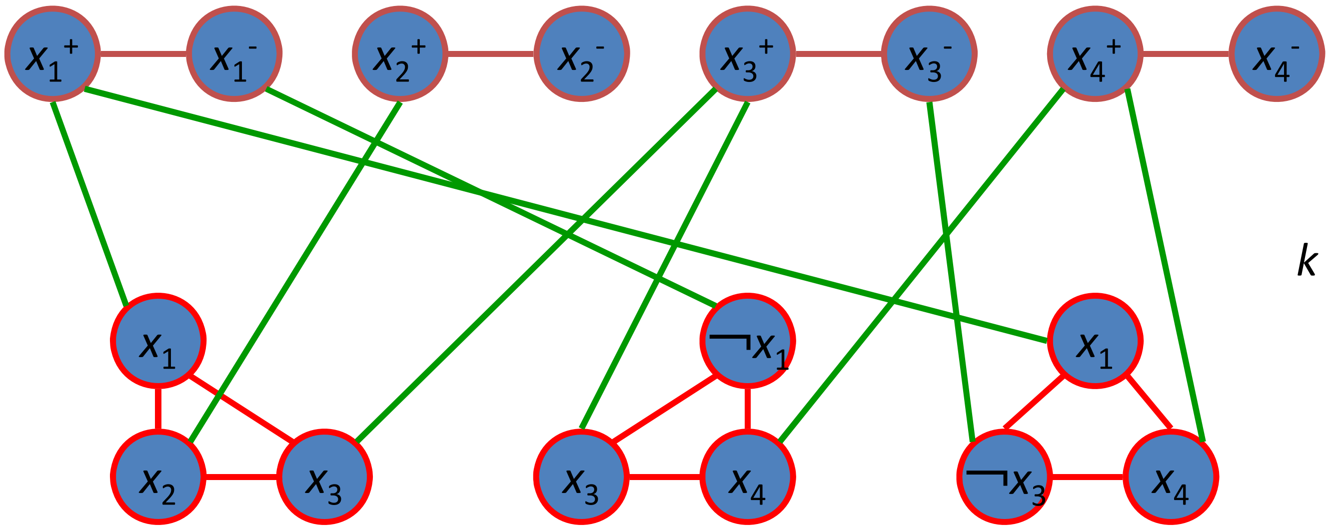
2. よって被覆 S はx_i^+ と x_i^- からちょうど一つと，各項C_jからちょうど二つのリテラルを含む

3. つまり各項 C_j はSに含まれないリテラルl_iをちょうど一つだけ含み，そこにつながる辺は変数頂点で被覆されている．

⇒ 以下の条件を満たす割当ては F を充足する：$\left\{ \begin{array}{l} x_i^+ \in S \text{ なら} x_i=1 \\ x_i^- \in S \text{ なら } x_i=0 \end{array} \right\}$

例：　$F(x_1,x_2,x_3,x_4) = (x_1 \vee x_2 \vee x_3) \wedge (\neg x_1 \vee x_3 \vee x_4) \wedge (x_1 \vee \neg x_3 \vee x_4)$

G



$k = 4+2 \times 3=10$

Q.E.D.

Theorem VC is NP-complete

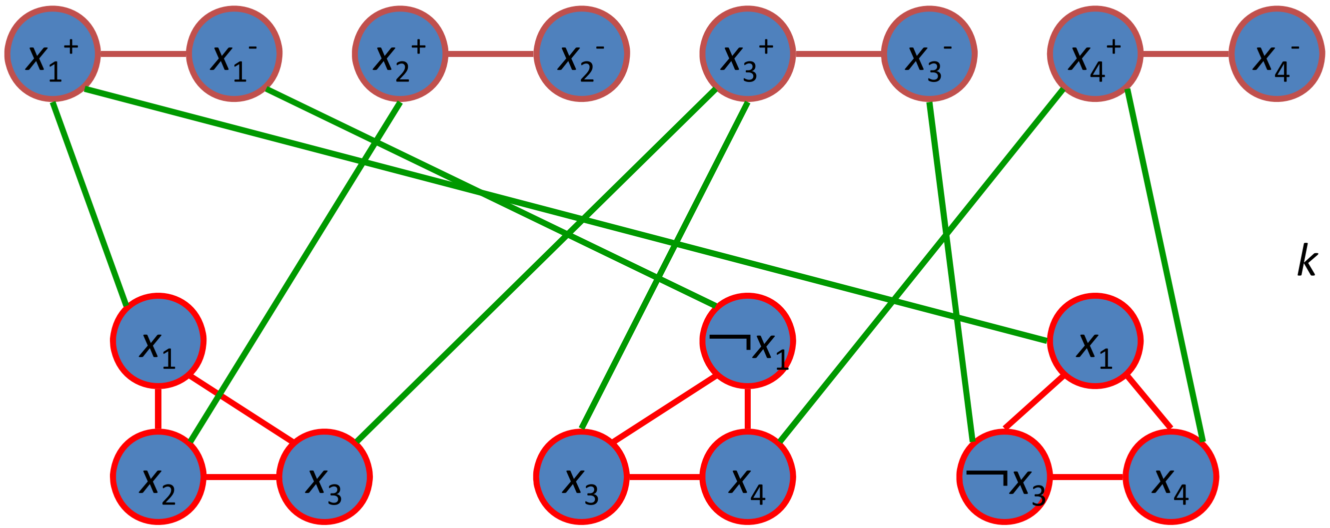
If G has a vertex cover of size k, there is an assignment that makes $F()=1$

1. From Observation, a cover S contains $2m$ vertices from the clauses, and n vertices from the variables.
2. Thus the cover S contains exactly one of x_i^+ and x_i^- and exactly two literals of a clause C_j.
3. Hence each clause C_j contains exactly one literal l_i which is not in S, and hence incident edge should be covered by a variable vertex.

$\Rightarrow$ The following assignment satisfies F: $\begin{cases} x_i=1 \text{ if } x_i^+ \text{ in } S \\ x_i=0 \text{ if } x_i^- \text{ in } S \end{cases}$

Ex: $F(x_1,x_2,x_3,x_4) = (x_1 \vee x_2 \vee x_3) \wedge (\neg x_1 \vee x_3 \vee x_4) \wedge (x_1 \vee \neg x_3 \vee x_4)$

G



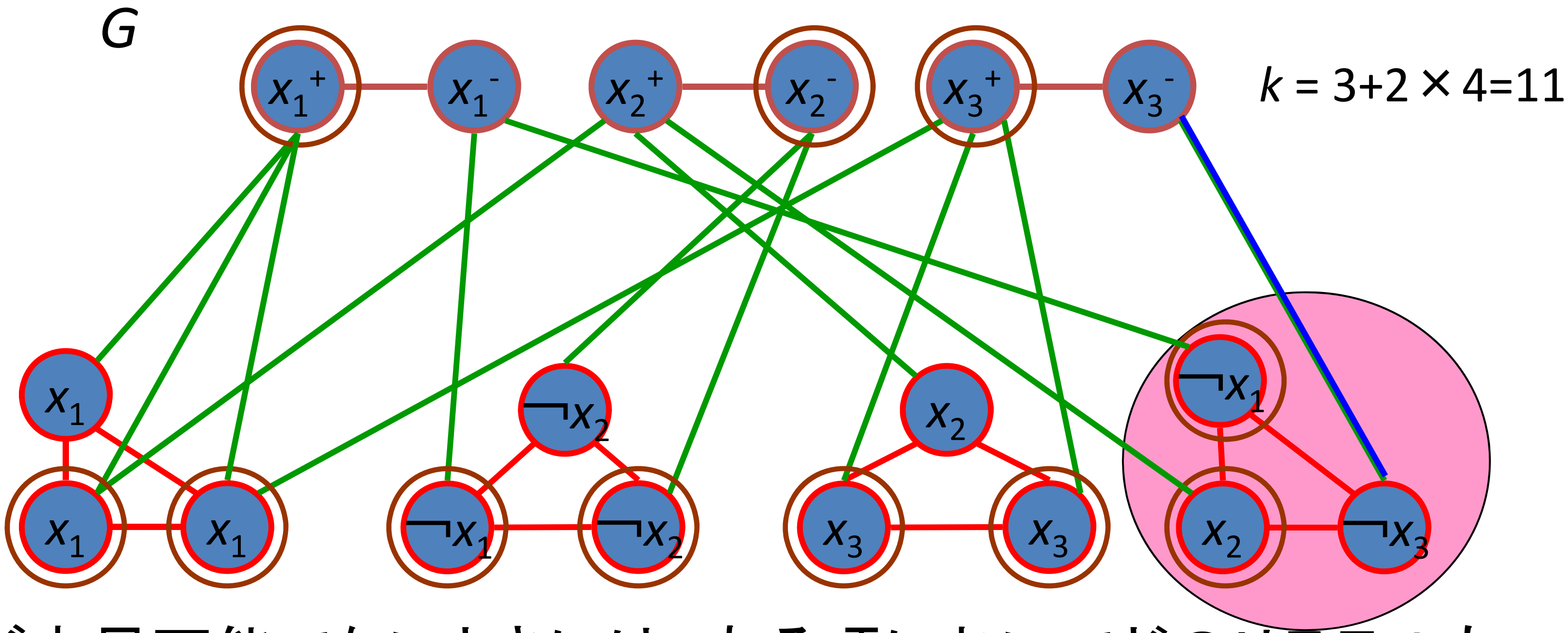
$k = 4+2 \times 3=10$

Q.E.D.

定理 VCはNP完全問題… 補足

命題論理式が充足可能でないときにはどうなるのか?

$F(x_1,x_2,x_3) = (x_1 \lor x_1 \lor x_1) \land (\neg x_1 \lor \neg x_2 \lor \neg x_2) \land (x_2 \lor x_3 \lor x_3) \land (\neg x_1 \lor x_2 \lor \neg x_3)$

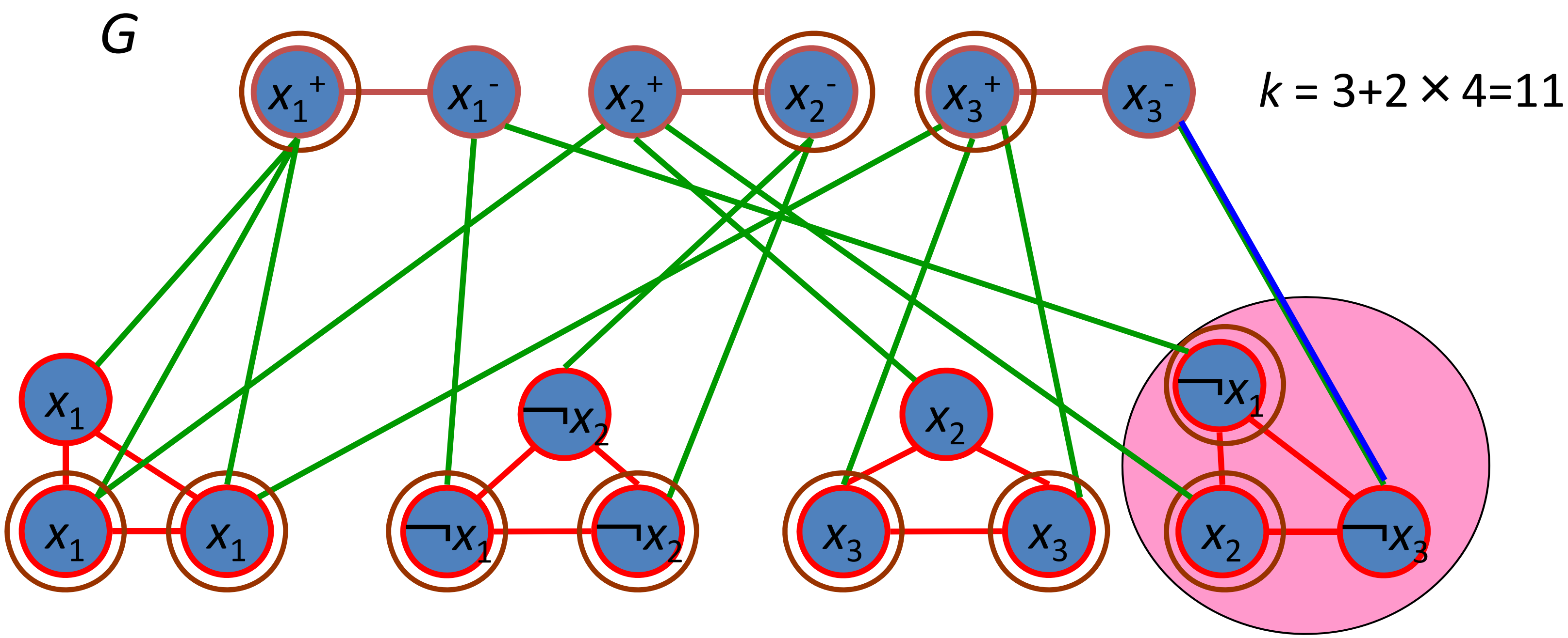


F が充足可能でないときには，ある項においてどのリテラルも変数側の頂点によって被覆されない．よって頂点被覆集合はこの項のリテラルを三つともふくまなければならない．
したがって頂点被覆集合は少なくとも大きさk+1になる．

Theorem VC is NP-complete… Addition

What happen if the formula is not satisfiable?

$$F(x_1,x_2,x_3) = (x_1 \vee x_1 \vee x_1) \wedge (\neg x_1 \vee \neg x_2 \vee \neg x_2) \wedge (x_2 \vee x_3 \vee x_3) \wedge (\neg x_1 \vee x_2 \vee \neg x_3)$$



If *F* is unsatisfiable, it contains at least one clause s. t. each literal is *not* covered by a vertex. So, Vertex Cover should contain *three* literals in the clause. Hence any vertex cover has size at least *k*+1.

6. 多項式時間計算可能性の解析

6.2. 完全性

定理

次数: 頂点につながる辺の本数

DHAM はグラフの最大次数が5でも NP完全

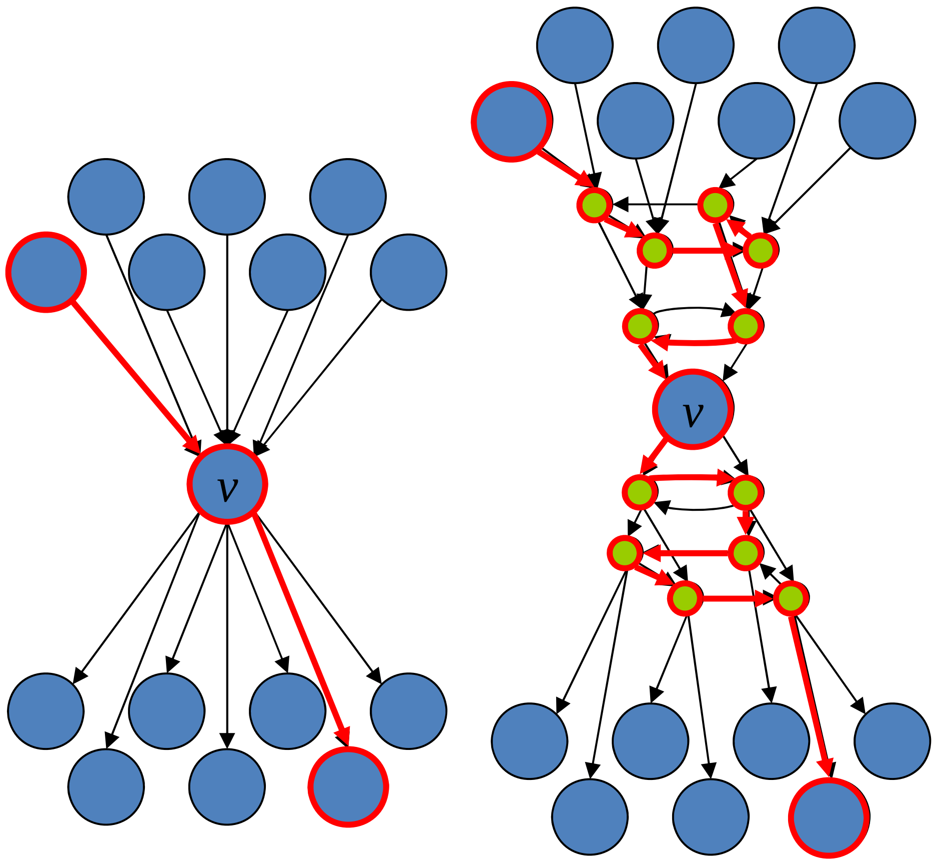
[証明]

DHAM ∈ NPなので $\mathrm{DHAM}_{\leqq 5}$ ∈NP.
よってDHAM $\leq_m^P \mathrm{DHAM}_{\leqq 5}$を示す

アイデア

「vに入る辺」や「vから出る辺」を
しかるべきガジェットで置き換える

元のグラフでvを通る
ハミルトン閉路は,
置き換えたグラフでvを
通るハミルトン閉路に
対応づけられる.



6. Analysis on Polynomial-Time Computability

6.2. Completeness

degree: the number of edges incident to a vertex

Theorem

DHAM is NP-complete even if maximum degree=5.

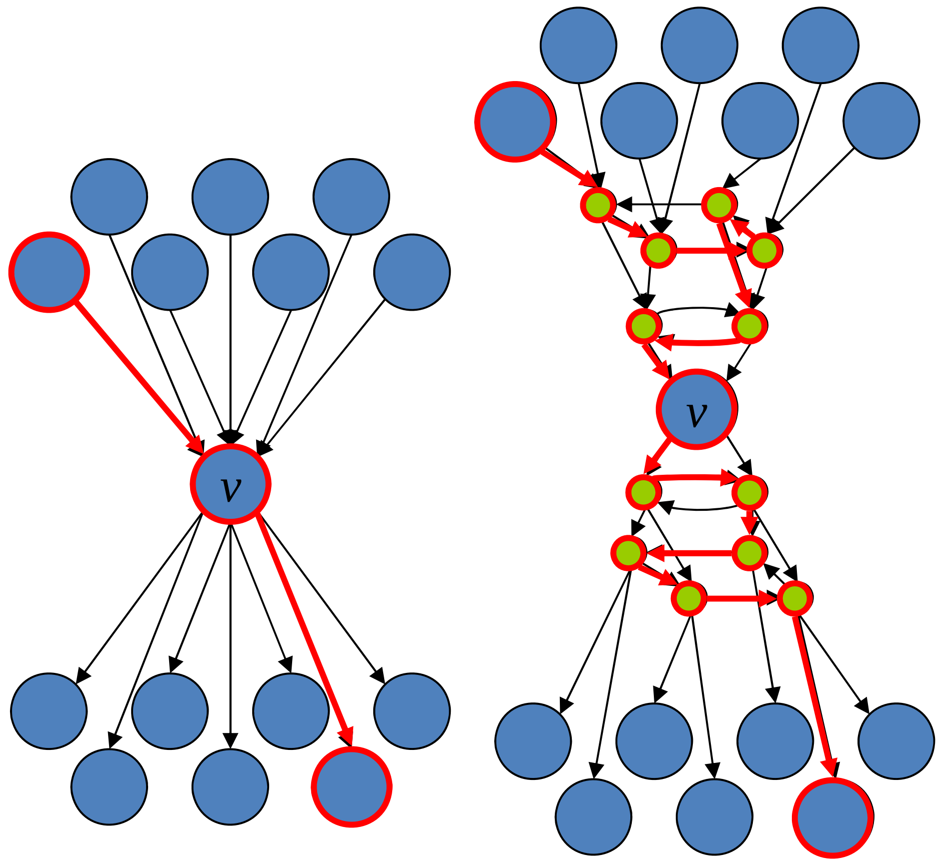
[Proof]

Since DHAM $\in$ NP, $\mathrm{DHAM}_{\leqq 5}$ $\in$ NP.
We show DHAM $\le_m^P$ $\mathrm{DHAM}_{\leqq 5}$.

Idea:

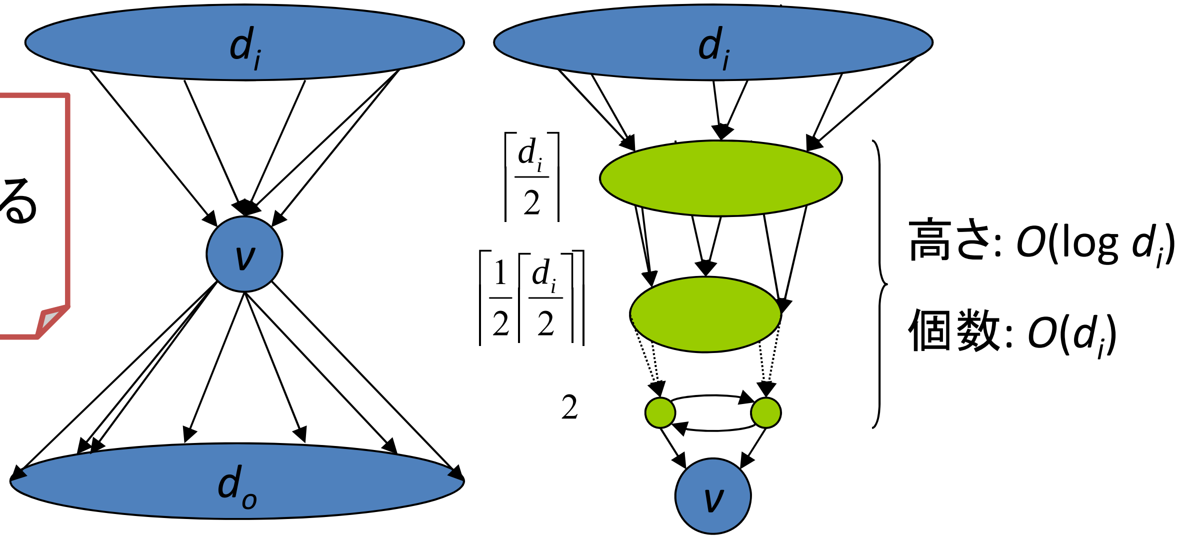
Replace the set of "arcs to v" and the set of "arcs from v" by a right 'gadget'.

A Hamiltonian cycle through *v* on the original graph corresponds to the Hamiltonian cycle through *v* on the resultant graph.



6.2.完全性

定理　DHAM はグラフの最大次数が5でも NP完全



ポイント:
- 上から下に閉路を通る
- 各頂点の次数≦5

[証明 (概要)]

次数≧6の各頂点 v に対して，v の頂点の周りの辺をガジェットで置き換える

1. もとのグラフ G の頂点数をn，辺数を m とすると，還元後に得られるグラフ G' の頂点数は $O(n+m)$ で辺数は $O(m)$ となる．よってこの還元はn と m の多項式時間で実行できる．
2. G' の各頂点の次数はたかだか 5.
3. G がハミルトン閉路をもつ ⇔ G' がハミルトン閉路をもつ

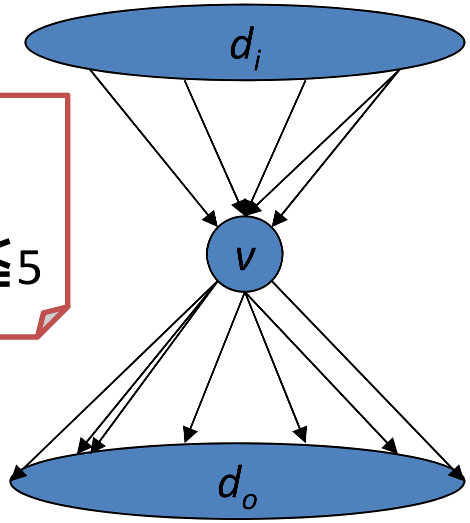
QED.

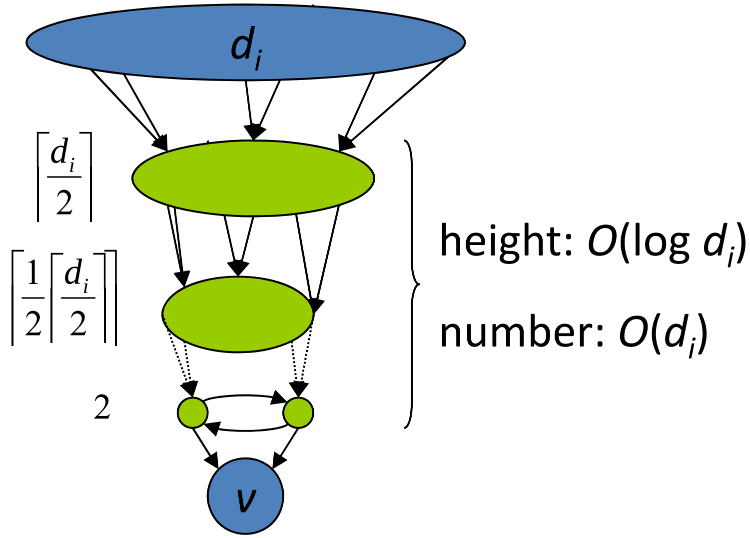
6.2. Completeness

Theorem DHAM is NP-complete even if max. degree=5.

Points:

- Up to down via cycle
- Each vertex has deg≦5





[Proof (sketch)]

For each vertex v of degree≧6, replace the edges around v by the gadget.

1. If the original graph G has n vertices with m edges, the resultant graph G' contains $O(n+m)$ vertices with $O(m)$ edges. Hence the reduction can be done in polynomial time of n & m.
2. Each vertex in G' has degree *at most* 5.
3. G has a Hamiltonian cycle ⇔ G' has a Hamiltonian cycle.

QED.

Addition (おまけ)

Many natural hard problems are either
- Poly-time solvable, or
- NP-hard

- R. Uehara, S. Iwata:
  Generalized Hi-Q is NP-complete,
  The Transactions of the IEICE, E73, p.270-273, 1990.
- P. Zhang, H. Sheng, R. Uehara:
  A Double Classification Tree Search Algorithm for
  Index SNP Selection, *BMC Bioinformatics*, 5:89, 2004.

- R. Uehara, S. Teramoto:
  Computational Complexity of a Pop-up Book,
  4th International Conference on Origami in Science, Mathematics, and Education, 2006.
- E. Demaine, M. Demaine, R. Uehara, T. UNO, Y. UNO:
  UNO is hard, even for a single player,
  Theoretical Computer Science, Vol. 521, pp.51-61, 2014.

- E. D. Demaine, Y. Okamoto, R. Uehara, and Y. Uno:
  Computational complexity and an integer programming model of Shakashaka,
  IEICE Trans. Vol. E97-A, No. 6, pp. 1213-1219, 2014.

残りの予定 (Schedule)

- 5/31(Thu): 多項式時間還元性に基づく完全性
  - 講義アンケート(アナウンスのみ)
  - レポート返却
- 5/31(Thu)チュートリアルアワー: 試験(Final Exam)
  - I1講義室で実施(13:30-15:10)
  - 30分過ぎたら退出可
  - 30分以上遅刻したら入室不可
  - 30点満点
  - スライドのコピー/手書きノート/筆記用具のみ(Copy of slides, hand-written note, and pens/pencils)
- 6月5日は，上原出張のため不在です．