コーダルグラフのサンドイッチ列挙

清見礼[1]　来嶋秀治[2]

Abstract

コーダルグラフは大きさが 4 以上の誘導部分サイクルをもたないグラフとして定義される．グラフ $\overline{G}$ と $\overline{G}$ の部分グラフ $\underline{G}$ が与えられた時, $\overline{G}$ の真部分グラフでかつ $\underline{G}$ を真部分グラフとして持つようなコーダルグラフ G をみつける問題は, コーダルグラフサンドイッチ問題と呼ばれ, NP-困難であることが知られている．我々は, $\overline{G}$ または $\underline{G}$ がコーダルグラフであれば, コーダルグラフサンドイッチ問題は容易に解くことが可能であることを示した．さらに, これを用いて $\overline{G}$ および $\underline{G}$ のいずれか一方がコーダルである場合に, $\overline{G}$ の部分グラフであり, $\underline{G}$ を部分グラフとして持つようなコーダルグラフをすべて列挙する問題のアルゴリズムを開発した．この結果は, 著者らの以前の結果 [5] の自然な拡張になっている．

An Algorithm for Chordal Graph Sandwich Enumeration Problem

Masashi Kiyomi[3]　Shuji Kijima[4]

Abstract

A graph is chordal iff its vertices do not induce any chordless cycle of length more than three. Given a graph $\overline{G}$ and $\underline{G}$ (where $\underline{G}$ is a subgraph of $\overline{G}$), the problem to find a graph G such that G is a proper subgraph of $\overline{G}$ and $\underline{G}$ is a proper subgraph of G is known as "chordal graph sandwich problem" and is NP-hard. We show that if either $\overline{G}$ or $\underline{G}$ is chordal, we can solve the chordal graph sandwich prolem in polynomial time. Besides, we develop an algorithm for enumerating every graph G that is a subgraph of $\overline{G}$ and a super graph of $\underline{G}$, under the condition that at least one of $\overline{G}$ and $\underline{G}$ is chordal. This result is a natural extension of our previous work [5].

1　はじめに

列挙とは, 領域 $\mathcal{D}$ と $\mathcal{D}$ 上で定義された条件 C が与えられたとき, C を満たすような $\mathcal{D}$ の元をもれなくかつ重複なく出力することである．例えば 4 頂点 1,2,3,4 からなる完全グラフの部分グラフにおいて, その完全グラフの全域木であるものの列挙は図 1 のようになる．与えられた条件を満たすようなグラフ構造の列挙は

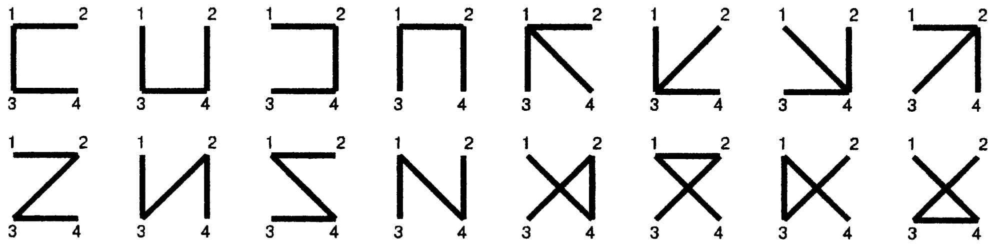

図 1: 4 頂点 1,2,3,4 からなる完全グラフ上で全域木であるものの列挙

データマイニング, 統計, ゲノム科学等のさまざまな分野への応用がある．

コーダルグラフは大きさが 4 以上の誘導部分サイクルをもたないグラフとして定義される．コーダルグラフはパーフェクトグラフの部分クラスであり, 様々な問題がこのクラス上で高速に解ける [2]．また, 隣接行列が, 掃き出しを行った際に fill in が発生しない行列の非零パターンであるなど, 応用上重要な性質を持つ．

[1]国立情報学研究所 masashi@grad.nii.ac.jp
[2]東京大学 kijima@misojiro.t.u-tokyo.ac.jp, JSPS Research Fellow
[3]National Institute of Informatics masashi@grad.nii.ac.jp
[4]University of Tokyo kijima@misojiro.t.u-tokyo.ac.jp, JSPS Research Fellow

グラフ $\overline{G}$ と $\overline{G}$ の部分グラフ $\underline{G}$ が与えられた時, $\overline{G}$ の真部分グラフでかつ $\underline{G}$ を真部分グラフとして持つようなコーダルグラフ G をみつける問題は, コーダルグラフサンドイッチ問題と呼ばれ, NP-困難であることが知られている [3]. 我々は, $\overline{G}$ または $\underline{G}$ がコーダルグラフであれば, コーダルグラフサンドイッチ問題は容易に解くことが可能であることを示した. さらにこれを用いて, $\overline{G}$ と $\underline{G}$ のいずれか一方がコーダルグラフである場合に, $\overline{G}$ の部分グラフであり, $\underline{G}$ を部分グラフとして持つようなコーダルグラフをすべて列挙する問題のアルゴリズムを開発した. この結果は, 著者らの以前の結果 [5] の自然な拡張になっている.

以下では, 2 節で本稿で用いるグラフに関する用語を説明し, 3 節でコーダルグラフの定義および諸性質について述べる. 4 節ではコーダルグラフサンドイッチ問題において, 入力が特殊な場合の多項式時間解法について述べる. さらに, 5 節ではこの結果を用いたコーダルグラフサンドイッチ列挙アルゴリズムについて述べ, 6 節でまとめを行う.

2 用語

頂点集合 V と, $V \times V$ 上の枝集合 E の組 (V, E) でグラフを表現する. グラフ G の頂点集合を $V(G)$, 枝集合を $E(G)$ で表す. 2 つのグラフ G_1, G_2 について, $V(G_1) \subseteq V(G_2), E(G_1) \subset E(G_2)$ が成り立つとき, $G_1 \subset G_2$ と書く. グラフ $G = (V, E)$ と V の部分集合 V' が与えられたとき,

$$G' = (V', \{e = \{v_1, v_2\} \in E \mid v_1 \in V', v_2 \in V'\})$$

を V' により G から誘導されるグラフという. また, G' は G の誘導部分グラフであるという. グラフ $G = (V, E)$ と枝 $e = \{v_1, v_2\} \in V \times V \setminus E$ が与えられたとき, グラフ $(V, E \cup \{e\})$ のことを $G + e$ と書く. 同様に, グラフ $G = (V, E)$ と枝 $e = \{v_1, v_2\} \in E$ が与えられたとき, グラフ $(V, E \setminus \{e\})$ のことを $G - e$ と書く.

3 コーダルグラフの諸性質

3.1 定義

定義 3.1 グラフ G が大きさが 4 以上の誘導部分サイクルをもたないとき, G はコーダルグラフである, という.

定義から明らかなように, コーダルグラフの任意の誘導部分グラフもまたコーダルグラフである.

3.2 単体的頂点

定義 3.2 グラフ G において, 頂点 v に隣接する頂点がクリークをなすとき, v は G の単体的頂点である, という.

コーダルグラフについては, 以下の定理が知られている.

定理 3.3 空でないコーダルグラフ G は単体的頂点をもつ [2].

3.3 完全消去列

コーダルグラフの大きな特徴として, コーダルグラフの頂点集合は完全消去列をもつ, というものが挙げられる. 完全消去列の定義は以下である.

定義 3.4 n 頂点からなるグラフ G が与えられたとき, G の頂点集合の列 $(v_1, \ldots, v_n)$ で, 以下の性質を満たすものを, G の頂点の完全消去列という.

v_i $(1 \leq i \leq n)$ は, 頂点 $v_i, \ldots, v_n$ によって G から誘導されるグラフの単体的頂点である.

なお, 以後では混乱がない場合, G の頂点の完全消去列のことを G の完全消去列と呼ぶことがある. 完全消去列はコーダルグラフの特徴付けを行う. すなわち, 以下の定理が成り立つ.

定理 3.5 グラフ G がコーダルグラフであることと, グラフ G が完全消去列をもつことは同値である [2].

また, 完全消去列については, 以下の定理も知られている.

定理 3.6 コーダルグラフ G と G の任意の頂点 v が与えられたとき, v が最後尾になるような G の頂点の完全消去列が存在する [2].

さらに, コーダルグラフ G が与えられたとき, 完全消去列 P を G のサイズの線形の時間で求めることが可能である.

定理 3.7 G が与えられたとき, G に完全消去列が存在するかどうかの判定は G の頂点数および枝数の和の線形の時間で行うことができる. さらに, もし G に完全消去列が存在するならば, G の完全消去列 P を同様の時間で求めることが可能である.

この定理により, 与えられたグラフがコーダルグラフかどうかの判定も, 入力の線形の時間で行うことができることに注意する.

4 コーダルグラフサンドイッチ問題

入力 $\overline{G}$ と $\underline{G}$ がともに一般の単純グラフである場合, コーダルグラフサンドイッチ問題は NP-困難であり [3], 解を高速に見つける一般的なアルゴリズムを開発することは絶望的である. そこで, 本章では, $\overline{G}$ と $\underline{G}$ のいずれかがコーダルグラフである場合について考える. 入力グラフにこのような制約を加えることにより, 入力の多項式時間でグラフサンドイッチ問題の解を見つけることが可能になる. 以下に, 本章で扱う定理を述べる.

定理 4.1 与えられたグラフ $\overline{G}$ とグラフ $\underline{G}$ がコーダルグラフで, $\underline{G} \subset \overline{G}$ が成り立ち, かつ $|E(\underline{G})|+1 < |E(\overline{G})|$ である場合,

$$\underline{G} \subset G \subset \overline{G}$$

を満たすコーダルグラフ G が存在する.

以下では, この定理を示すため, 以下のより強い定理を扱う.

定理 4.2 与えられたグラフ $\overline{G}$ とグラフ $\underline{G}$ がコーダルグラフで, $\underline{G} \subset \overline{G}$ が成り立ち, かつ $|E(\underline{G})|+1 < |E(\overline{G})|$ である場合,

$$\underline{G} + e^* \subset \overline{G},\ \underline{G} + e^* \text{はコーダルグラフ}$$

を満たす枝 $e^* \in E(\overline{G}) \setminus E(\underline{G})$ が存在する.

Proof 一般に，コーダルグラフ G に孤立点 v を追加したグラフ $(V(G)+v, E(G))$ はコーダルグラフであるので，$V(\overline{G}) = V(\underline{G})$ であるとして一般性を失わない．ここでは $V(\overline{G}) = V(\underline{G}) = \{v_1, v_2, \ldots, v_n\}$ とする．$\overline{G}$ の枝で，$\underline{G}$ の枝ではないものの集合を $\widetilde{E}$ とおく．コーダルグラフ $\underline{G}$ の完全消去列 $P = (p_1, p_2, \ldots, p_n)$ を考える．$\underline{G}_i$ $(1 \leq i \leq n)$ で，頂点 $p_i, \ldots, p_n$ によって $\underline{G}$ から誘導されるグラフを表すものとする．また，$\overline{G}_i$ $(1 \leq i \leq n)$ で，頂点 $p_i, \ldots, p_n$ によって $\overline{G}$ から誘導されるグラフを表すものとする．$\widetilde{E}_i = E(\overline{G}_i) \setminus E(\underline{G}_i)$ $(i \in \{1, 2, \ldots, n\})$ とする．完全消去列 P 中の p_i 以降の点を両端点とするような $\widetilde{E}$ の枝が存在するような i の最大値を s とする．すなわち，

$$s = \max\{i \in \{1, 2, \ldots, n\} \mid \widetilde{E}_i \neq \emptyset\}$$

である．ここで，s の定義より，明らかにすべての $\widetilde{E}_s$ の枝は端点として s をもつことを注意しておく（つまり，s 以外の頂点が 2 本以上の $\widetilde{E}_s$ の枝に接続することはない）．いま，$|\widetilde{E}_s| = 1$ または $|\widetilde{E}_s| > 1$ である．以下で，それぞれの場合に対して，定理の主張が成り立つことを示す．

$|\widetilde{E}_s| \quad 1$ の場合

$\widetilde{E}_s$ の（唯一つの）元を $\tilde{e}$ とする．e^* として $\tilde{e}$ を選べば，$\underline{G} + e^*$ の完全消去列が存在し，ゆえに $\underline{G} + e^*$ がコーダルグラフであることを示す．

P が $\underline{G}$ の完全消去列であることから，p_i $(1 \leq i < s)$ は $\underline{G}_i$ の単体的頂点である．つまり，$\underline{G}_i$ において p_i に接する頂点はクリークをなす．$\tilde{e}$ は p_s と p_j $(j > s)$ を結ぶ枝であるので，p_i $(1 \leq i < s)$ と接続することはない．ゆえに，$\underline{G}_i + \tilde{e}$ において p_i に接する頂点集合は，$\underline{G}$ において p_i に接する頂点集合と等しく，$\underline{G} + \tilde{e}$ においてもクリークをなす．よって，p_i $(1 \leq i < s)$ は $\underline{G}_i + \tilde{e}$ の単体的頂点である．

コーダルグラフの定義より，コーダルグラフの誘導部分グラフはコーダルグラフである．いま，$|\widetilde{E}_s| = 1$ であるので，$\underline{G}_s + \tilde{e} = \overline{G}_s$ である．よって，$\underline{G}_s + \tilde{e}$ はコーダルグラフ $\overline{G}$ の誘導部分グラフであり，ゆえにコーダルグラフである．そこで，$\underline{G}_s + \tilde{e}$ には完全消去列 $(\bar{p}_s, \ldots, \bar{p}_n)$ が存在する．

以上より，$(p_1, \ldots, p_{s-1}, \bar{p}_s, \ldots, \bar{p}_n)$ は $\underline{G} + \tilde{e}$ の完全消去列である．ゆえに，$e^* = \tilde{e}$ とすれば $\underline{G} + e^*$ はコーダルグラフである．

$|\widetilde{E}_s| > 1$ の場合

$\underline{G}_s$ はコーダルグラフであるので，s を最後尾とするような完全消去列 $(p'_s, \ldots, p'_{n-1}, p'_n \quad p_s)$ をもつ．そこで，$p'_i = p_i$ $(1 \leq i < s)$ とおけば，$P' = (p'_1, \ldots, p'_n)$ は $\underline{G}$ の完全消去列である．$\underline{G}'_i$ $(1 \leq i \leq n)$ で，頂点 $p'_i, \ldots, p'_n$ によって $\underline{G}$ から誘導されるグラフを表すものとする．また，$\overline{G}'_i$ $(1 \leq i \leq n)$ で，頂点 $p'_i, \ldots, p'_n$ によって $\overline{G}$ から誘導されるグラフを表すものとする．$\widetilde{E}'_i = E(\overline{G}'_i) \setminus E(\underline{G}'_i)$ $(i \in \{1, 2, \ldots, n\})$ とする．

$$s' = \max\{i \in \{1, 2, \ldots, n\} \mid \widetilde{E}'_i \neq \emptyset\}$$

とする．s' は完全消去列 P において s より後に登場する頂点であるので，$\widetilde{E}'_{s'} \subseteq \widetilde{E}_s$ である．よって，s' は $\widetilde{E}'_{s'}$ の枝高々1 本にしか接続することはできない．つまり，$|\widetilde{E}'_{s'}| = 1$ である．よって，$\widetilde{E}'_{s'}$ の（唯一つの）元を e^* とすれば，$|\widetilde{E}_s| = 1$ の場合と全く同様に，$\underline{G} + e^*$ はコーダルグラフである．□

さらに，上の定理から以下の定理を簡単に導くことができる．

定理 4.3 与えられたグラフ $\overline{G}$ とグラフ $\underline{G}$ がコーダルグラフで，$\underline{G} \subset \overline{G}$ が成り立ち，かつ $E(\underline{G}) < E(\overline{G}) - 1$ である場合，

$$\underline{G} \subset \overline{G} - e^*,\ \overline{G} - e^* \text{はコーダルグラフ}$$

を満たす枝 $e^* \in E(\overline{G}) \setminus E(\underline{G})$ が存在する．

Proof $E(\underline{G}) < E(\overline{G}) - 2$ である場合，定理 4.2 を再帰的に用いる（$\underline{G}$ を $\underline{G} + e^*$ と再帰的におきなおしていく）ことで，$E(\underline{G}) = E(\overline{G}) - 2$ とすることができる．$E(\underline{G}) = E(\overline{G}) - 2$ の場合，$\underline{G} + e^*$ は $\overline{G}$ から $\widetilde{E}$ の枝を 1 本取り除いたコーダルグラフである．□

以上で, $\overline{G}$ および $\underline{G}$ がコーダルグラフであった場合, コーダルグラフサンドイッチ問題が容易に解けることが分かった. この結果は, $\overline{G}$ または $\underline{G}$ の一方のみがコーダルグラフであった場合に拡張することが可能である.

定理 4.4 与えられたグラフ $\overline{G}$ とグラフ $\underline{G}$ のいずれか一方がコーダルグラフである場合,

$$\underline{G} \subset G \subset \overline{G}$$

を満たすコーダルグラフ G が存在するかどうかの判定は, 入力の多項式時間で行うことができる.

Proof 以下では, $\underline{G}$ がコーダルグラフである場合について証明する. $\overline{G}$ がコーダルグラフである場合も, まったく同様にして題意を示すことができる.

条件を満たすような G が存在するとする. すると, 定理 4.2 より,

$$\underline{G} + e^* \subseteq G \subset \overline{G},\ \underline{G} + e^* \text{はコーダルグラフ}$$

を満たす枝 $e^* \in E(G) \setminus E(\underline{G})$ が必ず存在する. いま, $E(G) \subset E(\overline{G})$ であるので, $E(\overline{G}) \setminus E(\underline{G})$ のすべての枝 e について $\underline{G} + e$ がコーダルグラフになるかどうかを調べれば十分である. □

5 コーダルグラフサンドイッチ列挙

本節では, $\underline{G}$ と $\overline{G}$ を入力として, $\underline{G} \subseteq G \subseteq \overline{G}$ を満たす G を列挙するアルゴリズムについて考える. ここで, 入力 $\underline{G}$ または $\overline{G}$ のいずれか一方はコーダルグラフであるとする. なお, $\overline{G}$ がコーダルである場合と, $\underline{G}$ がコーダルである場合の列挙アルゴリズムのアイデアは全く同様なものである. 本稿では紙面の制約のため, $\overline{G}$ がコーダルである場合のみについて述べることとする. 本節で扱う列挙アルゴリズムは, Avis と福田 [1] により提案された逆探索と呼ばれる手法を用いている.

5.1 逆探索

逆探索は, 深さ優先探索の一種である. いま, 列挙したい対象の集合を $\mathcal{F}$ とする. まず, $\mathcal{F}$ のそれぞれの元について, その元の親となる元を定めることで $\mathcal{F}$ の元の間に親子関係と呼ばれる関係を定義する. この際, 親子関係が以下の性質を満たすように注意する.

- 根と呼ばれるいくつか (典型的には 1 つ) の元は親をもたない.
- 根以外の元は必ず丁度 1 つの親をもつ.
- 自分が自分の先祖にならない. つまり, ある元から, その元の親, またその元の親, ... と辿っていったとき, もとの元に戻ることはない.

このとき, 親子関係を $\mathcal{F}$ の元が頂点となるような有効グラフで表すと, グラフは全張森 (根が 1 つなら全張木) になる. 図 2 にグラフ表現の例を示す. すべての $\mathcal{F}$ の元からいずれかの根へと至るパスが一意に定まることが分かる.

それぞれの枝を逆向きに辿ることで, $\mathcal{F}$ のすべての元を丁度 1 度だけ訪れる木探索アルゴリズムを容易に考えることができる. ここで, この木探索アルゴリズムが効率がよいものであるためには, うまく親子関係を定義することにより, 以下の条件が満たされている必要がある.

- すべての根を容易に見つけることができる.
- すべての元に対して, 自分を親とする元すべてを効率よく見つけることができる.
- 子がいないすべての元に対して, 子がいないということを効率よく判定することができる.

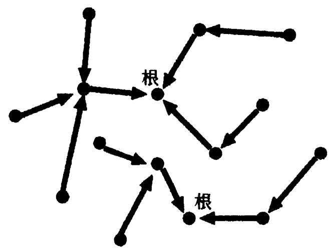

図 2: 列挙対象の全張森. 矢印の向きは子から親.

5.2 親子関係

コーダルグラフ $\overline{G}$ とグラフ $\underline{G}$ が与えられたとき, $\underline{G} \subseteq G \subseteq \overline{G}$ を満たすようなコーダルグラフ G の集合上での親子関係を以下のように定める.

- 根は $\{\overline{G}\}$ である.
- $G \neq \overline{G}$ の親は, G と $\overline{G}$ を入力としたとき定理 4.2 で得られる $G + e^*$ である.

ここで, 定理 4.2 の証明が, 構成的に e^* を得ていることに注意する. しかし, 一般にはコーダルグラフの完全消去列は複数存在するので, 定理 4.2 の証明での e^* の選び方には冗長性が生じてしまう. ただし, 定理 4.2 の証明に従って, e^* を求める冗長性のないアルゴリズム A を 1 つ固定することは容易であるので, 実際にはこうして e^* を求めるアルゴリズム A を 1 つ定めた上で, このアルゴリズムにより得られる $G + e^*$ を G の親と定義することにする.

5.3 アルゴリズム

以下に, コーダルグラフ $\overline{G}$ とグラフ $\underline{G}$ が与えられたとき, $\underline{G} \subseteq G \subseteq \overline{G}$ を満たすようなコーダルグラフ G を列挙するアルゴリズムを示す.

procedure enum_chordal(G, $\overline{G}$, $\underline{G}$)
G : **コーダルグラフ**, $\overline{G}$: **コーダルグラフ**, $\underline{G}$: **グラフ**;
begin
　G を出力;
　for $E(G) \setminus E(\underline{G})$ のすべての枝 e **do**
　　if $G - e$ はコーダルグラフ **and**
　　　$G - e$ の親は G **then**
　　　enum_chordal($G - e$, $\overline{G}$, $\underline{G}$);
　end for
end.

逆探索の説明より, $G = \overline{G}$ として, 上記のアルゴリズムを呼び出せば, $\underline{G} \subseteq G \subseteq \overline{G}$ を満たすようなコーダルグラフ G を列挙することができることが分かる.

5.4 計算量

前小節のアルゴリズムにおいて，コーダルグラフを 1 つ出力するのにかかる時間について考える．$\overline{G}$ の頂点数を n, 枝数を m とする．

コーダルグラフ G と G の枝 e が与えられたとき，$G-e$ がコーダルグラフであるかどうかの判定は $\mathrm{O}(m)$ でできる．また，$G-e$ がコーダルグラフであるとき，$G-e$ の親は $\mathrm{O}(m)$ で求めることが可能であるので，$G-e$ の親が G であるかどうかの判定にかかる時間は $\mathrm{O}(m)$ である．グラフ G が与えられたとき，$E(G)\setminus E(\underline{G})$ のすべての枝 e に対してこれらの処理を行うので，コーダルグラフ 1 つ当たりにかかる時間は

$$\mathrm{O}(m^2)$$

であることがわかる．

定理 5.1 コーダルグラフ $\overline{G}$ とグラフ $\underline{G}$ が与えられたとき，$\underline{G}\subseteq G\subseteq\overline{G}$ を満たすようなコーダルグラフ G を列挙するアルゴリズムで，出力 *1* つ当たりにかかる時間が $O(m^2)$ であるものが存在する．ただし，m は $\overline{G}$ の枝数である．

6 まとめ

コーダルグラフサンドイッチ問題について，入力グラフの片方がコーダルグラフであれば，多項式時間判定が可能であることを示した．さらに，グラフのサンドイッチ列挙というものを考え，コーダルグラフサンドイッチ列挙は，入力グラフの一方がコーダルグラフであれば，一つ当たり多項式時間で行うことが可能であることを示した．論文 [5] において，筆者らは部分グラフ列挙および supergraph 列挙について述べているが，本稿の結果はこれらの結果の拡張になっている．そこで，[5] においてコーダルグラフとともに考察されている，区間グラフについて，同様にグラフサンドイッチ問題およびサンドイッチ列挙を考えることができるが，本稿と同様な結果が得られるかどうかはまだ分かっていない．

参考文献

[1] Avis, D., and Fukuda, K.: Reverse search for enumeration, Discrete Applied Mathematics **65** (1996) 21–46

[2] Golumbic, M. C.: Algorithmic graph theory and perfect graphs (2nd), Elsevier, 2004

[3] Golumbic, M. C., Kaplan, H., and Shamir R.: Graph sandwich problems, Journal of Algorithms **19** (1995) 449–473

[4] Kiyomi, M., and Uno, T.: Generating chordal graphs included in given graphs, IEICE Transactions on Information and Systems **E89-D** (2006) 763–770

[5] Kiyomi, M., Kijima, S., and Uno, T.: Listing chordal graphs and interval graphs, Proceedings of 32nd International Workshop on Graph-Theoretic Concepts in Computer Science (LNCS になる予定)