Finding and counting given length cycles

N. Alon, R. Yuster, and U. Zwick

Algorithmica **17** (1997) 209–223.

Speaker: Joseph, Chuang-Chieh Lin
Supervisor: Professor Maw-Shang Chang

Computation Theory Laboratory
Department of Computer Science and Information Engineering
National Chung Cheng University, Taiwan

January 5, 2010

Outline

Outline

Problem

Given a graph and an integer k, decide whether a given graph $G = (V, E)$ contains a simple cycle of length k.

- This problem is **NP**-complete.
- However, for every fixed k, it can be solved in either $O(|V||E|)$ time or $O(|V|^\omega \log |V|)$ time ($\omega < 2.376$)
  - Monien. *Annuals of Discret. Math.*, 1985.
  - Alon, Yuster, Zwick. *J. ACM*, 1995.

Introduction

Problem

Given a graph and an integer k, decide whether a given graph $G = (V, E)$ contains a simple cycle of length k.

- This problem is **NP**-complete.
- However, for every fixed k, it can be solved in either $O(|V||E|)$ time or $O(|V|^\omega \log |V|)$ time ($\omega < 2.376$)
  - Monien. *Annuals of Discret. Math.*, 1985.
  - Alon, Yuster, Zwick. *J. ACM*, 1995.

The contributions of this paper

- An assortment of methods for finding and counting simple cycles of a given length in directed/undirected graphs.
- Most of the bounds depends solely on the number of edges of the input graph.
  - These bounds are of the form $O(|E|^{\alpha_k})$ or $O(|E|^{\beta_k} \cdot d(G)^{\gamma_k})$, where $\alpha_k, \beta_k, \gamma_k$ are some constants depending on k and $d(G)$ is the degeneracy of a graph (we will talk about it later).
- An application of color-coding.
  - Omitted in this talk due to insufficient time.

The contributions of this paper (contd.)

C_k: a simple cycle of length k; $d(G) \leq 2|E|^{1/2}$.

- ▷ In directed or undirected graphs:
  - A C_k in a directed or undirected graph $G = (V, E)$, if one exists, can be found in
    - $O(|E|^{2-2/k})$ time if k is even;
    - $O(|E|^{2-2/(k+1)})$ time if k is odd.
  - * A C_3 (triangle) can be found in $O(|E|^{2\omega/(\omega+1)}) = O(|E|^{1.41})$ time.
- ▷ In directed or undirected graphs (with the parameter $d(G)$):
  - A C_{4k-2} can be found in $O(|E|^{2-(1/2k)} \cdot d(G)^{1-1/k})$ time.
  - A C_{4k-1} and a C_{4k} can be found in $O(|E|^{2-1/k} \cdot d(G))$ time;
  - A C_{4k+1} can be found in $O(|E|^{2-1/k} \cdot d(G)^{1+1/k})$ time;

The contributions of this paper (contd.)

- In an undirected graph, finding even cycles is even faster:
  - A C_{4k-2} (if one exists) can be found in $O(|E|^{2-(1/2k)(1+1/k)})$ time.
  - A C_{4k} can be found in $O(|E|^{2-(1/k-1/(2k+1))})$ time;
  - * A C_4 can be found in $O(|E|^{4/3})$ time; and a C_6 can be found in $O(|E|^{13/8})$ time.

The contributions of this paper (contd.)

Cycle	Complexity	Cycle	Complexity
C_3	$\|E\|^{1.41}$, $\|E\| \cdot d(G)$	C_7	$\|E\|^{1.75}$, $\|E\|^{3/2} \cdot d(G)$
C_4	$\|E\|^{1.5}$, $\|E\| \cdot d(G)$	C_8	$\|E\|^{1.75}$, $\|E\|^{3/2} \cdot d(G)$
C_5	$\|E\|^{1.67}$, $\|E\| \cdot d(G)^2$	C_9	$\|E\|^{1.8}$, $\|E\|^{3/2} \cdot d(G)^{3/2}$
C_6	$\|E\|^{1.67}$, $\|E\|^{3/2} \cdot d(G)^{1/2}$	C_{10}	$\|E\|^{1.8}$, $\|E\|^{5/3} \cdot d(G)^{2/3}$

Table: Finding small cycles in *directed* graphs – some of the new results in this paper.

Cycle	Complexity	Cycle	Complexity
C_4	$\|E\|^{1.34}$	C_8	$\|E\|^{1.7}$
C_6	$\|E\|^{1.63}$	C_{10}	$\|E\|^{1.78}$

Table: Finding small cycles in *undirected* graphs – some of the new results in this paper.

Outline

q-representative

- A p-set is a set of size p.

Definition

Let $\mathcal{F}$ be a collection of p-sets. A subcollection $\hat{\mathcal{F}} \subseteq \mathcal{F}$ is q-representative for $\mathcal{F}$ if:

- *for every q-set B, there exists a set $A \in \mathcal{F}$ such that $A \cap B = \emptyset$ if and only if there exists a set $A' \in \hat{\mathcal{F}}$ such that $A' \cap B = \emptyset$.*

For example, let

$$\mathcal{F} = \{\{1,2,3\},\{2,3,4\},\{1,3,6\},\{4,5,6\},\{4,5,7\}\}$$

be a collection of 3-sets.

- Choose $\hat{\mathcal{F}} = \{\{1,2,3\},\{4,5,7\}\}$.
  - $\hat{\mathcal{F}}$ is NOT 3-representative (conider $\{2,4,7\}$).
  - $\hat{\mathcal{F}}$ is NOT 2-representative (consider $\{1,5\}$).
- Choose $\hat{\mathcal{F}}' = \{\{1,2,3\},\{1,3,6\},\{4,5,7\}\}$.
  - $\hat{\mathcal{F}}'$ is NOT 3-representative (consider $\{1,5,8\}$).
- Choose $\hat{\mathcal{F}}'' = \{\{1,2,3\},\{2,3,4\},\{1,3,6\},\{4,5,7\}\}$.
  - $\hat{\mathcal{F}}''$ is still NOT 3-representative (consider $\{1,3,7\}$).

q-representative (contd.)

For example, let

$$\mathcal{F} = \{\{1,2,3\},\{2,3,4\},\{1,3,6\},\{4,5,6\},\{4,5,7\}\}$$

be a collection of 3-sets.

- Choose $\hat{\mathcal{F}} = \{\{1,2,3\},\{4,5,7\}\}$.
  - $\hat{\mathcal{F}}$ is NOT 3-representative (conider $\{2,4,7\}$).
  - $\hat{\mathcal{F}}$ is NOT 2-representative (consider $\{1,5\}$).
- Choose $\hat{\mathcal{F}}' = \{\{1,2,3\},\{1,3,6\},\{4,5,7\}\}$.
  - $\hat{\mathcal{F}}'$ is NOT 3-representative (consider $\{1,5,8\}$).
- Choose $\hat{\mathcal{F}}'' = \{\{1,2,3\},\{2,3,4\},\{1,3,6\},\{4,5,7\}\}$.
  - $\hat{\mathcal{F}}''$ is still NOT 3-representative (consider $\{1,3,7\}$).

For example, let

$$\mathcal{F} = \{\{1,2,3\},\{2,3,4\},\{1,3,6\},\{4,5,6\},\{4,5,7\}\}$$

be a collection of 3-sets.

- Choose $\hat{\mathcal{F}} = \{\{1,2,3\},\{4,5,7\}\}$.
  - $\hat{\mathcal{F}}$ is NOT 3-representative (conider $\{2,4,7\}$).
  - $\hat{\mathcal{F}}$ is NOT 2-representative (consider $\{1,5\}$).
- Choose $\hat{\mathcal{F}}' = \{\{1,2,3\},\{1,3,6\},\{4,5,7\}\}$.
  - $\hat{\mathcal{F}}'$ is NOT 3-representative (consider $\{1,5,8\}$).
- Choose $\hat{\mathcal{F}}'' = \{\{1,2,3\},\{2,3,4\},\{1,3,6\},\{4,5,7\}\}$.
  - $\hat{\mathcal{F}}''$ is still NOT 3-representative (consider $\{1,3,7\}$).

q-representative (contd.)

Consider another example. Let

$$\mathcal{F} = \{\{a, b, c\}, \{b, c, d\}, \{c, d, e\}, \{d, e, f\}, \{e, f, a\}, \{f, a, b\}\}$$

be a collection of 3-sets.

- Choose $\hat{\mathcal{F}} = \{\{a, b, c\}, \{d, e, f\}\}$.
  - $\hat{\mathcal{F}}$ is r-representative for any integer $r \geq 1$.

Two important results

Lemma 2.1 (Bollobás. 1965)

Any collection $\mathcal{F}$ of p-sets, no matter how large it is, has a q-representative subcollection of size at most $\binom{p+q}{p}$.

Theorem 2.2 (Monien. 1985)

Given a collection $\mathcal{F}$ of p-sets. There is an $O(pq \cdot \sum_{i=0}^{q} p^i \cdot |\mathcal{F}|)$ time algorithm to find a q-representative subcollection $\mathcal{F}' \subseteq \mathcal{F}$ where $|\mathcal{F}'| \leq \sum_{i=0}^{q} p^i$.

The key lemma

Lemma 2.3

$\mathcal{F}$: a collection of p-sets; $\mathcal{G}$: a collection of q-sets (p, q are fixed). We can either (1) find $A \in \mathcal{F}, B \in \mathcal{G}$ s.t. $A \cap B = \emptyset$ or (2) decide that no such two sets exist in $O(|\mathcal{F}| + |\mathcal{G}|)$ time.

Proof.

- Use Monien's algorithm to find (in $O(|\mathcal{F}| + |\mathcal{G}|)$ time):
  - a q-representative $\hat{\mathcal{F}} \subseteq \mathcal{F}$ s.t. $|\hat{\mathcal{F}}| \leq \sum_{i=0}^{q} p^i$,
  - a p-representative $\hat{\mathcal{G}} \subseteq \mathcal{G}$ s.t. $|\hat{\mathcal{G}}| \leq \sum_{i=0}^{p} q^i$.
- **Claim:** If $\exists A \in \mathcal{F}$, $\exists B \in \mathcal{G}$ such that $A \cap B = \emptyset$, then $\exists A' \in \hat{\mathcal{F}}$, $\exists B' \in \hat{\mathcal{G}}$ such that $A' \cap B' = \emptyset$.
  - ⋆ if $A \cap B = \emptyset$, by the definition, $\exists A' \in \hat{\mathcal{F}}$ such that $A' \cap B = \emptyset$ (similarly, $\exists B' \in \hat{\mathcal{G}}$ such that $A' \cap B' = \emptyset$).
- After finding $\hat{\mathcal{F}}$ and $\hat{\mathcal{G}}$, it is enough to check whether they contain two disjoint sets in constant time.

The key lemma

Lemma 2.3

$\mathcal{F}$: a collection of p-sets; $\mathcal{G}$: a collection of q-sets (p, q are fixed). We can either (1) find $A \in \mathcal{F}, B \in \mathcal{G}$ s.t. $A \cap B = \emptyset$ or (2) decide that no such two sets exist in $O(|\mathcal{F}| + |\mathcal{G}|)$ time.

Proof.

- Use Monien's algorithm to find (in $O(|\mathcal{F}| + |\mathcal{G}|)$ time):
  - a q-representative $\hat{\mathcal{F}} \subseteq \mathcal{F}$ s.t. $|\hat{\mathcal{F}}| \leq \sum_{i=0}^{q} p^i$,
  - a p-representative $\hat{\mathcal{G}} \subseteq \mathcal{G}$ s.t. $|\hat{\mathcal{G}}| \leq \sum_{i=0}^{p} q^i$.
- **Claim:** If $\exists A \in \mathcal{F}$, $\exists B \in \mathcal{G}$ such that $A \cap B = \emptyset$, then $\exists A' \in \hat{\mathcal{F}}$, $\exists B' \in \hat{\mathcal{G}}$ such that $A' \cap B' = \emptyset$.
  - ⋆ if $A \cap B = \emptyset$, by the definition, $\exists A' \in \hat{\mathcal{F}}$ such that $A' \cap B = \emptyset$ (similarly, $\exists B' \in \hat{\mathcal{G}}$ such that $A' \cap B' = \emptyset$).
- After finding $\hat{\mathcal{F}}$ and $\hat{\mathcal{G}}$, it is enough to check whether they contain two disjoint sets in constant time.

The key lemma

Lemma 2.3

$\mathcal{F}$: a collection of p-sets; $\mathcal{G}$: a collection of q-sets (p, q are fixed). We can either (1) find $A \in \mathcal{F}, B \in \mathcal{G}$ s.t. $A \cap B = \emptyset$ or (2) decide that no such two sets exist in $O(|\mathcal{F}| + |\mathcal{G}|)$ time.

Proof.

- Use Monien's algorithm to find (in $O(|\mathcal{F}| + |\mathcal{G}|)$ time):
  - a q-representative $\hat{\mathcal{F}} \subseteq \mathcal{F}$ s.t. $|\hat{\mathcal{F}}| \leq \sum_{i=0}^{q} p^i$,
  - a p-representative $\hat{\mathcal{G}} \subseteq \mathcal{G}$ s.t. $|\hat{\mathcal{G}}| \leq \sum_{i=0}^{p} q^i$.
- **Claim:** If $\exists A \in \mathcal{F}$, $\exists B \in \mathcal{G}$ such that $A \cap B = \emptyset$, then $\exists A' \in \hat{\mathcal{F}}$, $\exists B' \in \hat{\mathcal{G}}$ such that $A' \cap B' = \emptyset$.
  - ⋆ if $A \cap B = \emptyset$, by the definition, $\exists A' \in \hat{\mathcal{F}}$ such that $A' \cap B = \emptyset$ (similarly, $\exists B' \in \hat{\mathcal{G}}$ such that $A' \cap B' = \emptyset$).
- After finding $\hat{\mathcal{F}}$ and $\hat{\mathcal{G}}$, it is enough to check whether they contain two disjoint sets in constant time.

The key lemma

Lemma 2.3

$\mathcal{F}$: a collection of p-sets; $\mathcal{G}$: a collection of q-sets (p, q are fixed). We can either (1) find $A \in \mathcal{F}, B \in \mathcal{G}$ s.t. $A \cap B = \emptyset$ or (2) decide that no such two sets exist in $O(|\mathcal{F}| + |\mathcal{G}|)$ time.

Proof.

- Use Monien's algorithm to find (in $O(|\mathcal{F}| + |\mathcal{G}|)$ time):
  - a q-representative $\hat{\mathcal{F}} \subseteq \mathcal{F}$ s.t. $|\hat{\mathcal{F}}| \leq \sum_{i=0}^{q} p^i$,
  - a p-representative $\hat{\mathcal{G}} \subseteq \mathcal{G}$ s.t. $|\hat{\mathcal{G}}| \leq \sum_{i=0}^{p} q^i$.
- **Claim:** If $\exists A \in \mathcal{F}$, $\exists B \in \mathcal{G}$ such that $A \cap B = \emptyset$, then $\exists A' \in \hat{\mathcal{F}}$, $\exists B' \in \hat{\mathcal{G}}$ such that $A' \cap B' = \emptyset$.
  - ⋆ if $A \cap B = \emptyset$, by the definition, $\exists A' \in \hat{\mathcal{F}}$ such that $A' \cap B = \emptyset$ (similarly, $\exists B' \in \hat{\mathcal{G}}$ such that $A' \cap B' = \emptyset$).
- After finding $\hat{\mathcal{F}}$ and $\hat{\mathcal{G}}$, it is enough to check whether they contain two disjoint sets in constant time.

□

A useful lemma from Monien

Lemma 2.4 (Monien. 1985)

Let $G = (V, E)$ be a directed/undirected graph, let $v \in V$, and let $k \geq 3$. A C_k passing through v, if one exists, can be found in $O(|E|)$ time.

The algorithm of finding C_{2k} and C_{2k-1}

Theorem 2.5

Deciding whether a directed/undirected graph $G = (V, E)$ contains simple cycles of length exactly $2k - 1$ and of length exactly $2k$, and finding such cycles if it does, can be done in $O(|E|^{2-1/k})$ time.

Proof of Theorem 2.5

- Let $\Delta = |E|^{1/k}$.
- $v \in V$ is of high degree: $\deg(v) \geq \Delta$.
  - ⋆ G contains $\leq 2|E|/\Delta = O(|E|^{1-1/k})$ high-degree vertices.

▽ We describe an $O(|E|^{2-1/k})$ time algorithm for finding a C_{2k} in a directed graph $G = (V, E)$.

- ∗ The other cases are similar.

Sketch of the proof (algorithm):

I. Preprocessing (data reduction).

II. Find a C_{2k} containing u, v

$\Rightarrow$ Finding two paths: $u \xrightarrow{k} v$ & $v \xrightarrow{k} u$.

III. Make use of representative collections.

Proof of Theorem 2.5

- Let $\Delta = |E|^{1/k}$.
- $v \in V$ is of high degree: $\deg(v) \geq \Delta$.
  - ⋆ G contains $\leq 2|E|/\Delta = O(|E|^{1-1/k})$ high-degree vertices.

▽ We describe an $O(|E|^{2-1/k})$ time algorithm for finding a C_{2k} in a directed graph $G = (V, E)$.

- ∗ The other cases are similar.

Sketch of the proof (algorithm):

I. Preprocessing (data reduction).

II. Find a C_{2k} containing u, v

$\Rightarrow$ Finding two paths: $u \xrightarrow{k} v$ & $v \xrightarrow{k} u$.

III. Make use of representative collections.

Proof of Theorem 2.5

- Let $\Delta = |E|^{1/k}$.
- $v \in V$ is of high degree: $\deg(v) \geq \Delta$.
  - ⋆ G contains $\leq 2|E|/\Delta = O(|E|^{1-1/k})$ high-degree vertices.

▽ We describe an $O(|E|^{2-1/k})$ time algorithm for finding a C_{2k} in a directed graph $G = (V, E)$.

- ∗ The other cases are similar.

Sketch of the proof (algorithm):

I. Preprocessing (data reduction).

II. Find a C_{2k} containing u, v

$\Rightarrow$ Finding two paths: $u \xrightarrow{k} v$ & $v \xrightarrow{k} u$.

III. Make use of representative collections.

Actions similar to **data reductions**.

- Check any of these high-degree vertices lies on a C_{2k}.
  - If one of these vertices does lie on a C_{2k} then we are done.
  - Total time cost: $O(|E| \cdot |E|/\Delta) = O(|E|^{2-1/k})$.

- Otherwise, remove all the high-degree vertices and all edges incident to them, and then obtain a graph G'.
  - G' contains a $C_{2k} \Leftrightarrow G$ contains a C_{2k}.
  - $\max\{\deg_{G'}(v) \mid v \in V\} \leq \Delta = |E|^{1/k}$.
    - $\Rightarrow \leq |E| \cdot \Delta^{k-1} = |E|^{2-1/k}$ simple directed k-paths in G' (finding all of them: $O(|E|^{2-1/k})$ time).

Proof of Theorem 2.5 (contd.)

Actions similar to **data reductions**.

√ Check any of these high-degree vertices lies on a C_{2k}.

- If one of these vertices does lie on a C_{2k} then we are done.
- ⋆ Total time cost: $O(|E| \cdot |E|/\Delta) = O(|E|^{2-1/k})$.

√ Otherwise, remove all the high-degree vertices and all edges incident to them, and then obtain a graph G'.

- ⋆ G' contains a $C_{2k} \Leftrightarrow G$ contains a C_{2k}.
- ⋆ $\max\{\deg_{G'}(v) \mid v \in V\} \leq \Delta = |E|^{1/k}$.
  - $\Rightarrow \leq |E| \cdot \Delta^{k-1} = |E|^{2-1/k}$ simple directed k-paths in G' (finding all of them: $O(|E|^{2-1/k})$ time).

Actions similar to **data reductions**.

√ Check any of these high-degree vertices lies on a C_{2k}.

- If one of these vertices does lie on a C_{2k} then we are done.
- ⋆ Total time cost: $O(|E| \cdot |E|/\Delta) = O(|E|^{2-1/k})$.

√ Otherwise, remove all the high-degree vertices and all edges incident to them, and then obtain a graph G'.

- ⋆ G' contains a $C_{2k} \Leftrightarrow G$ contains a C_{2k}.
- ⋆ $\max\{\deg_{G'}(v) \mid v \in V\} \leq \Delta = |E|^{1/k}$.
  - $\Rightarrow$ $\leq |E| \cdot \Delta^{k-1} = |E|^{2-1/k}$ simple directed k-paths in G' (finding all of them: $O(|E|^{2-1/k})$ time).

Proof of Theorem 2.5 (contd.)

- $\surd$ Divide these paths into groups according to their endpoints.
  - ⋆ $O(|E|^{2-1/k})$ time & space by radix sort.
  - ■ We get a list $L = \{(u, v) \mid u \xrightarrow{k} v \text{ in } G'\}$.
- $\surd$ For each pair $(u, v) \in L$, get a collection $\mathcal{F}_{u,v}$ of $(k-1)$-sets.
  - ■ Each $(k-1)$-set in $\mathcal{F}_{u,v}$ corresponds to the $k-1$ intermediate vertices on some directed path $u \xrightarrow{k} v$.
- $\surd$ For each $(u, v) \in L$, check whether there exist two directed paths $u \xrightarrow{k} v$ and $v \xrightarrow{k} u$ that meet only at u, v.
  - ■ Such two paths exist if $\exists A \in \mathcal{F}_{u,v}, B \in \mathcal{F}_{v,u}$ s.t. $A \cap B = \emptyset$.
  - ⋆ Time cost: $O(|\mathcal{F}_{u,v}| + |\mathcal{F}_{v,u}|)$ (by the key lemma).
- ♠ The total time cost: $O(|E|^{2-1/k})$.
  - ⋆ Note that $\sum_{(u,v)\in L} |\mathcal{F}_{u,v}| = O(|E|^{2-1/k})$.

Proof of Theorem 2.5 (contd.)

√ Divide these paths into groups according to their endpoints.
- ⋆ $O(|E|^{2-1/k})$ time & space by radix sort.
- We get a list $L = \{(u, v) \mid u \xrightarrow{k} v \text{ in } G'\}$.

√ For each pair $(u, v) \in L$, get a collection $\mathcal{F}_{u,v}$ of $(k-1)$-sets.
- Each $(k-1)$-set in $\mathcal{F}_{u,v}$ corresponds to the $k-1$ intermediate vertices on some directed path $u \xrightarrow{k} v$.

√ For each $(u, v) \in L$, check whether there exist two directed paths $u \xrightarrow{k} v$ and $v \xrightarrow{k} u$ that meet only at u, v.
- Such two paths exist if $\exists A \in \mathcal{F}_{u,v}, B \in \mathcal{F}_{v,u}$ s.t. $A \cap B = \emptyset$.
- ⋆ Time cost: $O(|\mathcal{F}_{u,v}| + |\mathcal{F}_{v,u}|)$ (by the key lemma).

♠ The total time cost: $O(|E|^{2-1/k})$.
- ⋆ Note that $\sum_{(u,v)\in L} |\mathcal{F}_{u,v}| = O(|E|^{2-1/k})$.

Proof of Theorem 2.5 (contd.)

√ Divide these paths into groups according to their endpoints.

- ⋆ $O(|E|^{2-1/k})$ time & space by radix sort.
- We get a list $L = \{(u, v) \mid u \xrightarrow{k} v \text{ in } G'\}$.

√ For each pair $(u, v) \in L$, get a collection $\mathcal{F}_{u,v}$ of $(k-1)$-sets.

- Each $(k-1)$-set in $\mathcal{F}_{u,v}$ corresponds to the $k-1$ intermediate vertices on some directed path $u \xrightarrow{k} v$.

√ For each $(u, v) \in L$, check whether there exist two directed paths $u \xrightarrow{k} v$ and $v \xrightarrow{k} u$ that meet only at u, v.

- Such two paths exist if $\exists A \in \mathcal{F}_{u,v}, B \in \mathcal{F}_{v,u}$ s.t. $A \cap B = \emptyset$.
- ⋆ Time cost: $O(|\mathcal{F}_{u,v}| + |\mathcal{F}_{v,u}|)$ (by the key lemma).

♠ The total time cost: $O(|E|^{2-1/k})$.

- ⋆ Note that $\sum_{(u,v)\in L} |\mathcal{F}_{u,v}| = O(|E|^{2-1/k})$.

Proof of Theorem 2.5 (contd.)

- √ Divide these paths into groups according to their endpoints.
  - ⋆ $O(|E|^{2-1/k})$ time & space by radix sort.
  - ■ We get a list $L = \{(u, v) \mid u \xrightarrow{k} v \text{ in } G'\}$.
- √ For each pair $(u, v) \in L$, get a collection $\mathcal{F}_{u,v}$ of $(k-1)$-sets.
  - ■ Each $(k-1)$-set in $\mathcal{F}_{u,v}$ corresponds to the $k-1$ intermediate vertices on some directed path $u \xrightarrow{k} v$.
- √ For each $(u, v) \in L$, check whether there exist two directed paths $u \xrightarrow{k} v$ and $v \xrightarrow{k} u$ that meet only at u, v.
  - ■ Such two paths exist if $\exists A \in \mathcal{F}_{u,v}, B \in \mathcal{F}_{v,u}$ s.t. $A \cap B = \emptyset$.
  - ⋆ Time cost: $O(|\mathcal{F}_{u,v}| + |\mathcal{F}_{v,u}|)$ (by the key lemma).
- ♠ The total time cost: $O(|E|^{2-1/k})$.
  - ⋆ Note that $\sum_{(u,v)\in L} |\mathcal{F}_{u,v}| = O(|E|^{2-1/k})$.

A better result for triangles

Theorem 2.6

Deciding whether a directed/undirected graph $G = (V, E)$ contains a triangle, and finding one if it does, can be done in $O(E^{2\omega/(\omega+1)}) = O(E^{1.41})$ time.

Proof:

- Let $\Delta = |E|^{(\omega-1)/(\omega+1)}$.
- v is of *high-degree*: $\deg_G(v) > \Delta$ (low: otherwise).
  - ⋆ The number of high-degree vertices: $\leq 2|E|/\Delta$.

Proof of Theorem 2.6 (contd.)

- Consider all directed paths of length 2 in G whose intermediate vertex is of low degree.
  - ⋆ $\leq |E| \cdot \Delta$ such paths and can be found in $O(|E| \cdot \Delta)$ time.

- For each such 2-path $\{(u, v), (v, w)\}$, check whether u, v are connected by an edge (w, u).
  - No such a triangle is found $\Rightarrow$ triangles in G must be composed of three high-degree vertices.
  - Check whether there exists such a triangle using matrix multiplication ($O((|E|/\Delta)^\omega)$ time).

- Thus the total time cost is

$$O\left(|E| \cdot \Delta + \left(\frac{|E|}{\Delta}\right)^\omega\right) = O(|E|^{2\omega/(\omega+1)}).$$

Outline

Degeneracy (two equivalent definitions)

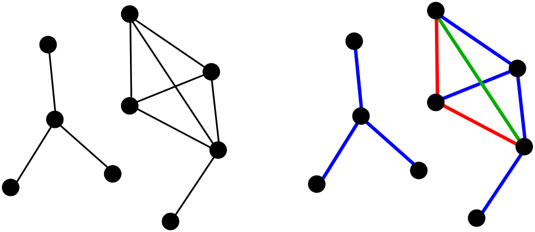
The degeneracy $d(G)$ of an undirected graph $G = (V, E)$ is:

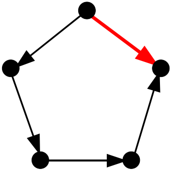
- ▷ the smallest number d for which there exists an *acyclic orientation* of G where all the out-degrees are at most d.
  - ∗ G is called d-degenerate.
- ▷ the maximum of the minimum degrees taken over all the subgraphs of G.

It's linearly related to arboricity of the graph.

- ▷ $a(G)$ is the minimum number of forests needed to cover all the edges of G.
- $a(G) \leq d(G) \leq 2 \cdot a(G) - 1$.
- It is easy to see that $a(G) \geq \lceil |E|/(|V| - 1) \rceil$.

Degeneracy (contd.)





Degeneracy (contd.)

Some examples.

- The degeneracy of any *tree* is 1.
- The degeneracy of any *cycle* is 2.
- The degeneracy of any *planar graph* is at most 5.
- For any graph $G = (V, E)$, we have $d(G) \leq 2|E|^{1/2}$ (when $|E| = \binom{|V|}{2}$, $V \approx (2|E|)^{1/2} < 2|E|^{1/2}$).
- If G is d-degenerate, then $|E| \leq d \cdot |V|$.

A folklore from Matula & Beck

Lemma 3.1 (Matula & Beck. *J. ACM*, 1983)

Let $G = (V, E)$ be a connected undirected graph. An acyclic orientation of G s.t. $\forall v \in V$, $d_{\mathrm{out}} \leq d(G)$ can be found in $O(|E|)$ time.

The other main result

Theorem 3.2

Let $G = (V, E)$ be a directed/undirected graph.

(i) *Deciding whether G contains a C_{4k-2}, and finding such a cycle if it does, can be done in $O(|E|^{2-1/k} \cdot d(G)^{1-1/k})$ time.*

(ii) *Deciding whether G contains a C_{4k-1} and a C_{4k}, and finding such cycles if it does, can be done in $O(|E|^{2-1/k} \cdot d(G))$ time.*

(iii) *Deciding whether G contains a C_{4k+1}, and finding such a cycle if it does, can be done in $O(|E|^{2-1/k} \cdot d(G)^{1+1/k})$ time.*

* If $d(G) \geq |E|^{1/(2k+1)}$, we can use the previous general algorithm.
  * $O(|E|^{2-1/(2k+1)}) \leq O(|E|^{2-1/k} \cdot d(G)^{1+1/k})$.
* Hence we assume that $d(G) \leq |E|^{1/(2k+1)}$.

The other main result

Theorem 3.2

Let $G = (V, E)$ be a directed/undirected graph.

(i) *Deciding whether G contains a C_{4k-2}, and finding such a cycle if it does, can be done in $O(|E|^{2-1/k} \cdot d(G)^{1-1/k})$ time.*

(ii) *Deciding whether G contains a C_{4k-1} and a C_{4k}, and finding such cycles if it does, can be done in $O(|E|^{2-1/k} \cdot d(G))$ time.*

(iii) *Deciding whether G contains a C_{4k+1}, and finding such a cycle if it does, can be done in $O(|E|^{2-1/k} \cdot d(G)^{1+1/k})$ time.*

- ∗ If $d(G) \geq |E|^{1/(2k+1)}$, we can use the previous general algorithm.
  - ⋆ $O(|E|^{2-1/(2k+1)}) \leq O(|E|^{2-1/k} \cdot d(G)^{1+1/k})$.
- ∗ Hence we assume that $d(G) \leq |E|^{1/(2k+1)}$.

Proof of Theorem 3.2

- Let $\Delta = |E|^{1/k}/d(G)^{1+1/k}$.
- $d(G) \leq |E|^{1/(2k+1)} = |E|^{\frac{1}{k}-(\frac{1}{2k+1})(\frac{k+1}{k})} \leq \Delta$.
- v has high-degree: $\deg(v) > \Delta$ (low-degree: otherwise).

√ Check whether $\exists$ a high-degree vertex lies on a C_{4k+1}.
  - ⋆ $O(|E|^2/\Delta)$ time.

√ If none of them lies on a C_{4k+1}, remove all the high-degree vertices from G, then obtain a graph $\tilde{G}$ with maximum degree $\leq \Delta$.

Proof of Theorem 3.2 (contd.)

- $d(\tilde{G}) \leq d(G)$.
  - * The degeneracy of a graph can only decrease when removing vertices and edges.

√ Get an acyclically oriented version G' of $\tilde{G}$ where each vertex has out-degree $\leq d(\tilde{G}) \leq d(G)$ (in $O(|E|)$ time).

- Consider the orientations, in G', of the edges on a $(2k+1)$-path in G.
  - * In at least one direction, $\exists \leq k$ *counterdirected* edges.

Proof of Theorem 3.2 (contd.)

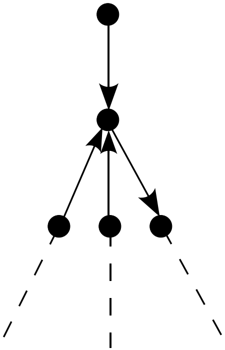


Fig.: Orientations of edges on paths.

Proof of Theorem 3.2 (contd.)

- The number of paths, **not necessarily directed**, of length $2k+1$ in $\tilde{G}$, is at most

$$
\begin{aligned}
& 2 \cdot 2|E| \cdot \sum_{i=0}^{k} \binom{2k}{i} \Delta^i d(G)^{2k-i} \\
= \quad & O\left(|E| \cdot k \binom{2k}{k} \cdot \sum_{i=0}^{k} \Delta^i d(G)^{2k-i} \right) \\
= \quad & O\left(|E| \cdot d(G)^{2k} \cdot \left(1 + \frac{\Delta}{d(G)} + \left(\frac{\Delta}{d(G)}\right)^2 + \ldots + \left(\frac{\Delta}{d(G)}\right)^k \right) \right) \\
= \quad & O\left(|E| \cdot d(G)^{2k} \cdot \frac{(\Delta/d(G))^{k+1} - 1}{\Delta/d(G) - 1} \right) \\
= \quad & O(|E| \Delta^k d(G)^k).
\end{aligned}
$$

- Similarly, the number of $2k$-paths in G is $O(|E|\Delta^k d(G)^{k-1})$.

Proof of Theorem 3.2 (contd.)

- By some further observations, we can lower the number of $2k+1$-paths and $2k$-paths a little bit.
  - ⋆ They are both $O(|E|\Delta^{k-1}d(G)^{k+1})$.
  - All the properly directed paths in G can be found in $O(|E|\Delta^{k-1}d(G)^{k+1})$ time.

√ Find a directed $(2k+1)$-path and a directed $2k$-path that close a directed simple cycle.
  - ⋆ $O(|E|\Delta^{k-1}d(G)^{k+1})$ time.

- The overall complexity:

$$O\left(\frac{|E|^2}{\Delta} + |E|\Delta^{k-1}d(G)^{k+1}\right) = O(|E|^{2-1/k}d(G)^{1+1/k}).$$

An application

Corollary 3.3

If a directed/undirected planar graph $G = (V, E)$ contains a C_5, then such a C_5 can be found in $O(|V|)$ time.

- ∗ $k = 1$ in this case.
- ∗ $|E| \leq d(G) \cdot |V|$ and $d(G) \leq 5$ ($\because$ G is planar).
- ⋆ $O(|E|^{2-1/k} \cdot d(G)^{1+1/k}) = O(|E| \cdot d(G)^2) = O(|V|)$.

Outline

Finding C_6 using color-coding

The following theorem follows by

- combining the previous ideas,
- using the $O(|E|^{2\omega/(\omega+1)})$ algorithm for finding triangles, and
- the color-coding method (Alon, Yuster, Zwick. *J. ACM*, 1995).

Theorem 3.4

Let $G = (V, E)$ be a directed/undirected graph. A C_6 in G, if one exists, can be found in either

- $O((|E| \cdot d(G))^{2\omega/(\omega+1)}) = O((|E| \cdot d(G))^{1.41})$ *expected time, or*
- $O((|E| \cdot d(G))^{1.41} \cdot \log |V|)$ *worst-case time.*

- Get an acyclically oriented G' of G with out-degree bounded by $d(G)$ (in $O(|E|)$ time).
- Suppose that G contains a C_6.

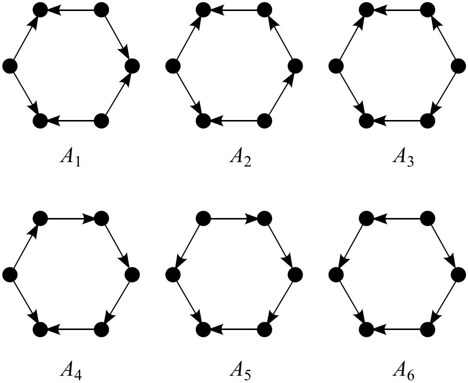
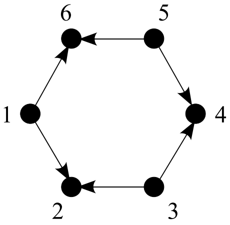


Fig.: Six possible orientations of a C_6 in G'

Color vertices of G' by six colors uniformly at random



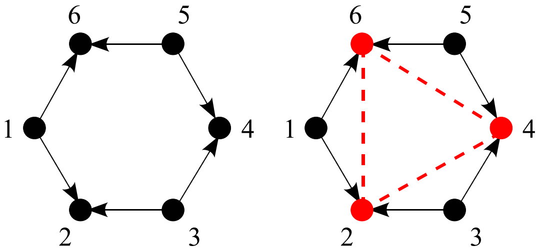
A_1

- Let A be a copy of A_1 in G'. A is well-colored if its vertices are consecutiely colored by 1 through 6.

⋆ $\mathbf{Pr}[A \text{ is well-colored}] = 6/6^6 = 1/6^5$.

- Assume that color 1 is assigned to a vertex having only out-going edges.

Create another undirected graph $G^* = (V^*, E^*)$ from G'



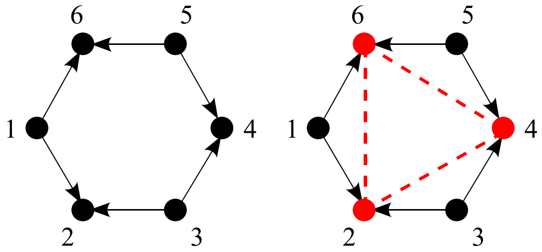
$c(v)$: the color number of v.

$V^* = \{v \in V \mid c(v) \in \{2, 4, 6\}\}$.

$$E^* = \{(u, v) \mid c(u) = 6, c(v) = 2, (\exists w \in V)(c(w) = 1, (w, u), (w, v) \in E')\}$$
$$\cup \{(u, v) \mid c(u) = 2, c(v) = 4, (\exists w \in V)(c(w) = 3, (w, u), (w, v) \in E')\}$$
$$\cup \{(u, v) \mid c(u) = 4, c(v) = 6, (\exists w \in V)(c(w) = 5, (w, u), (w, v) \in E')\}.$$

⋆ $|E^*| < |E| \cdot d(G)$.

Create another undirected graph $G^* = (V^*, E^*)$ from G'



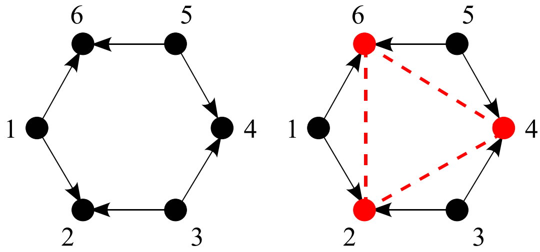
$c(v)$: the color number of v.

$$V^* = \{v \in V \mid c(v) \in \{2,4,6\}\}.$$

$$E^* = \{(u,v) \mid c(u) = 6, c(v) = 2, (\exists w \in V)(c(w) = 1, (w,u), (w,v) \in E')\}$$
$$\cup \{(u,v) \mid c(u) = 2, c(v) = 4, (\exists w \in V)(c(w) = 3, (w,u), (w,v) \in E')\}$$
$$\cup \{(u,v) \mid c(u) = 4, c(v) = 6, (\exists w \in V)(c(w) = 5, (w,u), (w,v) \in E')\}.$$

- ⋆ $|E^*| < |E| \cdot d(G)$.

Create another undirected graph $G^* = (V^*, E^*)$ from G'



$c(v)$: the color number of v.

$$V^* = \{v \in V \mid c(v) \in \{2, 4, 6\}\}.$$

$$E^* = \{(u, v) \mid c(u) = 6, c(v) = 2, (\exists w \in V)(c(w) = 1, (w, u), (w, v) \in E')\}$$
$$\cup \{(u, v) \mid c(u) = 2, c(v) = 4, (\exists w \in V)(c(w) = 3, (w, u), (w, v) \in E')\}$$
$$\cup \{(u, v) \mid c(u) = 4, c(v) = 6, (\exists w \in V)(c(w) = 5, (w, u), (w, v) \in E')\}.$$

- ⋆ $|E^*| < |E| \cdot d(G)$.

- $\exists$ an undirected triangle in G^* $\iff$ $\exists$ a well-colored A_1 in G'.

- Detecting triangles in G^*:
$O(|E^*|^{2\omega/(\omega+1)}) = O((|E| \cdot d(G))^{2\omega/(\omega+1)})$.

- Expected number of repetitions of the randomized coloring:
$6^5 = 7776$.

- The price for derandomization: $O(\log |V|)$.

Thank you!