Distinguishing Balls in Graphs

Jérémie Chalopin[1], Victor Chepoi[1], **Fionn Mc Inerney**[1], Sébastien Ratel[1], Yann Vaxès[1]

[1]LIS, Marseille, France

November 17, 2020

JGA 2020

Part I: Sample Compression Schemes

Samples

Given a domain U,

- a concept is a subset of U.
- a sample is a subset of U where each element has an associated label in $\{-1, +1\}$.

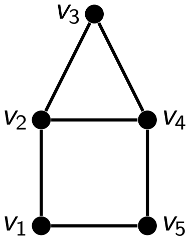
Realizable sample

For a concept class $\mathcal{C} \subseteq 2^U$ and a concept $C \in \mathcal{C}$, C realizes a sample X if, $\forall x \in X$ with label -1 and $\forall y \in X$ with label $+1$, $x \notin C$ and $y \in C$.

Samples

Given a domain U,

- a concept is a subset of U.
- a sample is a subset of U where each element has an associated label in $\{-1, +1\}$.



$\mathcal{C}$: balls of radius 1

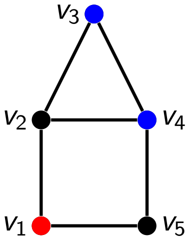
Realizable sample

For a concept class $\mathcal{C} \subseteq 2^U$ and a concept $C \in \mathcal{C}$, C realizes a sample X if, $\forall x \in X$ with label -1 and $\forall y \in X$ with label $+1$, $x \notin C$ and $y \in C$.

Samples

Given a domain U,

- a concept is a subset of U.
- a sample is a subset of U where each element has an associated label in $\{-1, +1\}$.



Realizable sample

For a concept class $\mathcal{C} \subseteq 2^U$ and a concept $C \in \mathcal{C}$, C realizes a sample X if, $\forall x \in X$ with label -1 and $\forall y \in X$ with label $+1$, $x \notin C$ and $y \in C$.

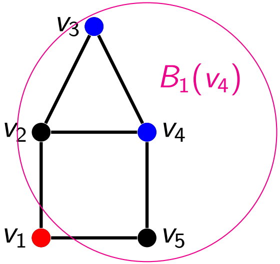
$\mathcal{C}$: balls of radius 1

$X_1 = \{v_1(-1), v_3(+1), v_4(+1)\}$

Samples

Given a domain U,

- a concept is a subset of U.
- a sample is a subset of U where each element has an associated label in $\{-1, +1\}$.



Realizable sample

For a concept class $\mathcal{C} \subseteq 2^U$ and a concept $C \in \mathcal{C}$, C realizes a sample X if, $\forall x \in X$ with label -1 and $\forall y \in X$ with label $+1$, $x \notin C$ and $y \in C$.

$\mathcal{C}$: balls of radius 1

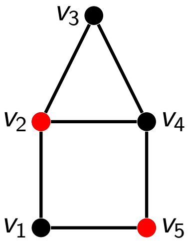
$X_1 = \{v_1(-1), v_3(+1), v_4(+1)\}$

$B_1(v_4)$ **realizes** X_1.

Samples

Given a domain U,

- a concept is a subset of U.
- a sample is a subset of U where each element has an associated label in $\{-1, +1\}$.



Realizable sample

For a concept class $\mathcal{C} \subseteq 2^U$ and a concept $C \in \mathcal{C}$, C realizes a sample X if, $\forall x \in X$ with label -1 and $\forall y \in X$ with label $+1$, $x \notin C$ and $y \in C$.

$\mathcal{C}$: balls of radius 1

$X_1 = \{v_1(-1), v_3(+1), v_4(+1)\}$

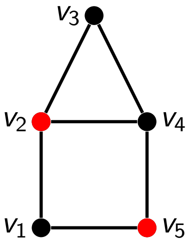
$B_1(v_4)$ **realizes** X_1.

$X_2 = \{v_2(-1), v_5(-1)\}$

Samples

Given a domain U,

- a concept is a subset of U.
- a sample is a subset of U where each element has an associated label in $\{-1, +1\}$.



Realizable sample

For a concept class $\mathcal{C} \subseteq 2^U$ and a concept $C \in \mathcal{C}$, C realizes a sample X if, $\forall x \in X$ with label -1 and $\forall y \in X$ with label $+1$, $x \notin C$ and $y \in C$.

$\mathcal{C}$: balls of radius 1

$X_1 = \{v_1(-1), v_3(+1), v_4(+1)\}$

$B_1(v_4)$ **realizes** X_1.

$X_2 = \{v_2(-1), v_5(-1)\}$

No ball **realizes** X_2.

(Labelled) Sample Compression Schemes

(Labelled) sample compression scheme of size k for a concept class $\mathcal{C}$

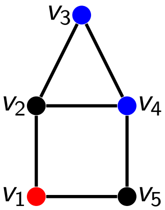
A function α **compressing** a realizable sample X into a (labelled) sub-sample $\alpha(X) \subseteq X$ of size at most k from which a function β **reconstructs** a concept $C \in \mathcal{C}$ realizing X.

(Labelled) Sample Compression Schemes

(Labelled) sample compression scheme of size k for a concept class $\mathcal{C}$

A function α **compressing** a realizable sample X into a (labelled) sub-sample $\alpha(X) \subseteq X$ of size at most k from which a function β **reconstructs** a concept $C \in \mathcal{C}$ realizing X.

$\mathcal{C}$: all balls of radius 1 in the house graph.



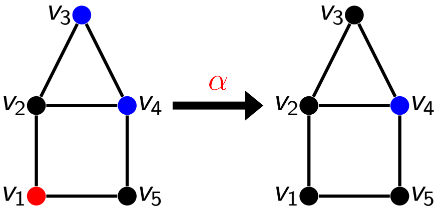
Sample X

(Labelled) Sample Compression Schemes

(Labelled) sample compression scheme of size k for a concept class $\mathcal{C}$

A function α **compressing** a realizable sample X into a (labelled) sub-sample $\alpha(X) \subseteq X$ of size at most k from which a function β **reconstructs** a concept $C \in \mathcal{C}$ realizing X.

$\mathcal{C}$: all balls of radius 1 in the house graph.



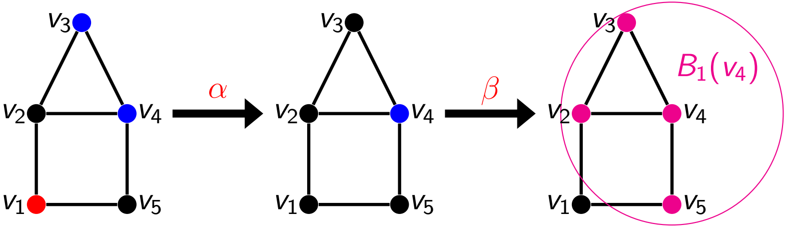
Sample X Sub-sample $\alpha(X)$

(Labelled) Sample Compression Schemes

(Labelled) sample compression scheme of size k for a concept class $\mathcal{C}$

A function α **compressing** a realizable sample X into a (labelled) sub-sample $\alpha(X) \subseteq X$ of size at most k from which a function β **reconstructs** a concept $C \in \mathcal{C}$ realizing X.

$\mathcal{C}$: all balls of radius 1 in the house graph.



Sample X Sub-sample $\alpha(X)$ Concept $B_1(v_4) \in \mathcal{C}$ realizing X

Related Work for Sample Compression Schemes (SCS)

- Defined for learning algorithms [Littlestone & Warmuth, 1986].

Related Work for Sample Compression Schemes (SCS)

- Defined for learning algorithms [Littlestone & Warmuth, 1986].
- Open problem in machine learning: does every concept class of VC-dimension d admit SCS of size $O(d)$? [Floyd & Warmuth, 1995].

Related Work for Sample Compression Schemes (SCS)

- Defined for learning algorithms [Littlestone & Warmuth, 1986].
- Open problem in machine learning: does every concept class of VC-dimension d admit SCS of size $O(d)$? [Floyd & Warmuth, 1995].
  - Labelled SCS (LSCS) of size $O(2^d)$ exist [Moran & Yehudayoff, 2016].
  - Ample sets admit LSCS of size d [Moran & Warmuth, 2016].
  - Unlabelled SCS (USCS) of size d exist for maximum families and were characterized for ample sets [Chalopin et al., 2019].
  - There's a family whose VC-dim $\leq d$ but **does not admit** USCS of size at most d [Pálvölgyi & Tardos, 2020].

Results Included in Part I

$\mathcal{B}(G)$: family of balls of any radius in a graph G.

$\mathcal{B}_r(G)$: family of balls of radius r in a graph G.

We show:

- USCS of size 2 for $\mathcal{B}(T)$ for trees T;
- LSCS (with extra information) of size 2 for $\mathcal{B}_r(T)$ for trees T;
- LSCS of size 3 for $\mathcal{B}(G)$ for interval graphs G.

USCS of Size 2 for $\mathcal{B}(T)$ for metric Trees T

For any realizable sample X,

Compressor $\alpha(X)$:

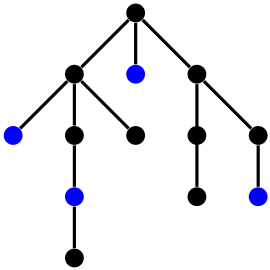
1. if $|X^+| \leq 1$, then $\alpha(X) = X^+$;
2. else $\alpha(X) = \{u^+, v^+\}$ (u^+, v^+: diametral pair of X^+).

USCS of Size 2 for $\mathcal{B}(T)$ for metric Trees T

For any realizable sample X,

Compressor $\alpha(X)$**:**

1. if $|X^+| \leq 1$, then $\alpha(X) = X^+$;
2. else $\alpha(X) = \{u^+, v^+\}$ (u^+, v^+: diametral pair of X^+).

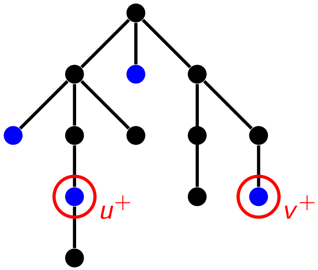


USCS of Size 2 for $\mathcal{B}(T)$ for metric Trees T

For any realizable sample X,

Compressor $\alpha(X)$**:**

1. if $|X^+| \leq 1$, then $\alpha(X) = X^+$;
2. else $\alpha(X) = \{u^+, v^+\}$ (u^+, v^+: diametral pair of X^+).



●: vertex in X^+

USCS of Size 2 for $\mathcal{B}(T)$ for metric Trees T

For any realizable sample X,

Compressor $\alpha(X)$**:**

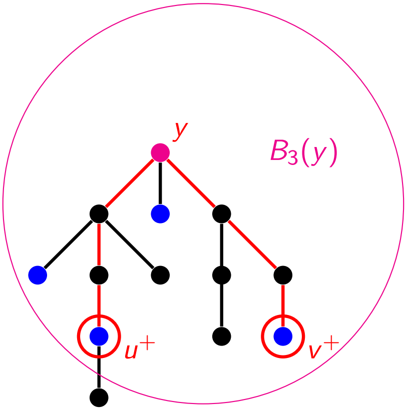
1. if $|X^+| \leq 1$, then $\alpha(X) = X^+$;
2. else $\alpha(X) = \{u^+, v^+\}$ (u^+, v^+: diametral pair of X^+).

Reconstructor $\beta(\alpha(X))$**:**

1. if $|\alpha(X)| = 0$, then $\beta(\alpha(X)) = \emptyset$;
2. else, if $|\alpha(X)| = 1$, then $\beta(\alpha(X)) = B_0(\alpha(X))$;
3. else $|\alpha(X)| = 2$ and $\beta(\alpha(X)) = B_r(y)$, where $r = \frac{1}{2}d(u^+, v^+)$ and y is the center of the path $P(u^+, v^+)$.

USCS of Size 2 for $\mathcal{B}(T)$ for metric Trees T

1. else $|\alpha(X)| = 2$ and $\beta(\alpha(X)) = B_r(y)$, where $r = \frac{1}{2}d(u^+, v^+)$ and y is the center of the path $P(u^+, v^+)$.

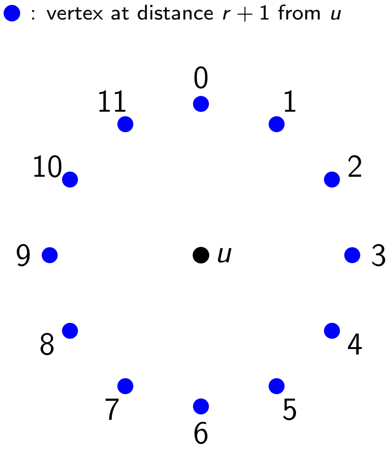


●: vertex in X^+

$\alpha(X) = \{u^+, v^+\}$

For $\mathcal{B}_r(T)$, Use Center Designators

If there exists $u \in X^-$ such that $B_r(y)$ realizes X and $d(u, y) = r + 1$, then there's a center designator.

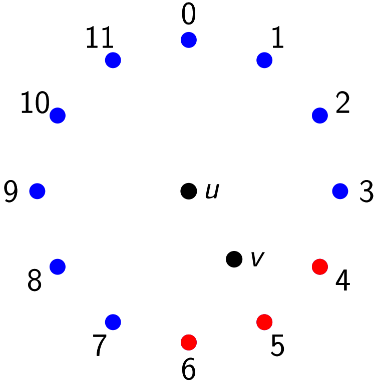


Blue vertices labelled by DFS from $u \in X^-$.

For $\mathcal{B}_r(T)$, Use Center Designators

If there exists $u \in X^-$ such that $B_r(y)$ realizes X and $d(u, y) = r + 1$, then there's a center designator.

● : vertex at distance $r+1$ from u

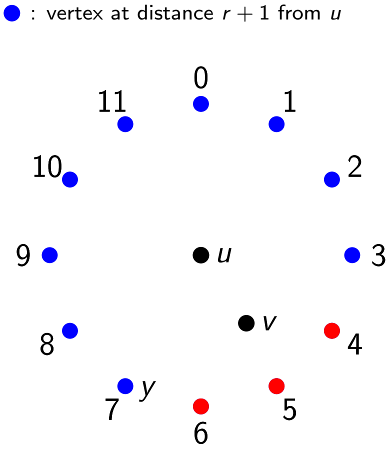


Blue vertices labelled by DFS from $u \in X^-$.

Any vertex $v \in X$ forbids a continuous interval of blue vertices.

For $\mathcal{B}_r(T)$, Use Center Designators

If there exists $u \in X^-$ such that $B_r(y)$ realizes X and $d(u, y) = r + 1$, then there's a center designator.



Blue vertices labelled by DFS from $u \in X^-$.

Any vertex $v \in X$ forbids a continuous interval of blue vertices.

If y follows a forbidden interval defined by $v \in X$ and $B_r(y)$ realizes X, then v is a center designator of u that designates y.

Two key lemmas for $\mathcal{B}_r(T)$

Lemma 1

Any ball containing a diametral pair of X^+, also contains X^+.

Lemma 2

Either there is a center designator or any ball containing a diametral pair of X^+ realizes X.

LSCS (w/ info) of Size 2 for $\mathcal{B}_r(T)$ for Trees T

Compressor $\alpha(X)$:

1. if $|X| \leq 1$, then $\alpha(X) = X$;
2. else, if $t \in X$ is a center designator of $s \in X^-$, then $\alpha(X) = \{t, s\}$;
3. otw, $\alpha(X) = \{u^+, v^+\}$ (diametral pair of X^+) or $\alpha(X) = X^+$ if $|X^+| = 1$.

LSCS (w/ info) of Size 2 for $\mathcal{B}_r(T)$ for Trees T

Compressor $\alpha(X)$:

1. if $|X| \leq 1$, then $\alpha(X) = X$;
2. else, if $t \in X$ is a center designator of $s \in X^-$, then $\alpha(X) = \{t, s\}$;
3. otw, $\alpha(X) = \{u^+, v^+\}$ (diametral pair of X^+) or $\alpha(X) = X^+$ if $|X^+| = 1$.

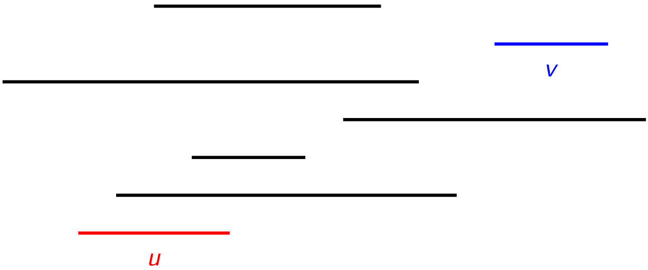
Reconstructor $\beta(\alpha(X))$:

1. if $|\alpha(X)| = |\alpha(X^-)| \leq 1$, then $\beta(\alpha(X))$ is any ball avoiding $\alpha(X^-)$;
2. else, if $|\alpha(X)| = |\alpha(X^+)| \geq 1$, then $\beta(\alpha(X))$ is any ball containing $\alpha(X^+)$;
3. otw, $\alpha(X) = \{t, s\}$ and $\beta(\alpha(X))$ is the ball $B_r(y)$ where t is the center designator of s that designates y.

LSCS of size 3 for $\mathcal{B}(G)$ for interval graphs G

Farthest pair of a subgraph H of an interval graph

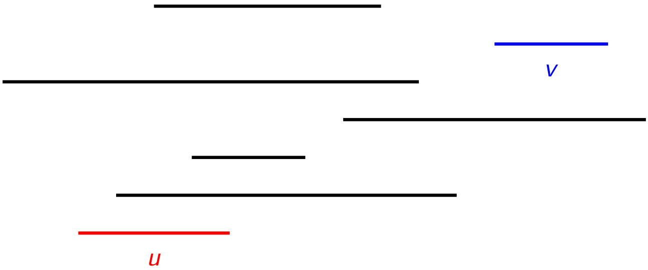
The vertices u, v in H such that the interval of u ends farthest to the left and the interval of v begins farthest to the right.



LSCS of size 3 for $\mathcal{B}(G)$ for interval graphs G

Farthest pair of a subgraph H of an interval graph

The vertices u, v in H such that the interval of u ends farthest to the left and the interval of v begins farthest to the right.



Key Lemma

Any ball containing a farthest pair of X^+, contains X^+.

Part II: Representation Maps

Representation Maps

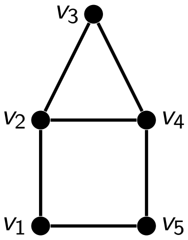
Representation map of size k for a concept class $\mathcal{C}$

A function ρ assigning to each $C \in \mathcal{C}$, a sample X_C of size at most k that is realizable by C, and such that, for any two $C_1, C_2 \in \mathcal{C}$, $C_1 \cap (\rho(C_1) \cup \rho(C_2)) \neq C_2 \cap (\rho(C_1) \cup \rho(C_2))$ if $C_1 \neq C_2$.

Representation Maps

Representation map of size k for a concept class $\mathcal{C}$

A function ρ assigning to each $C \in \mathcal{C}$, a sample X_C of size at most k that is realizable by C, and such that, for any two $C_1, C_2 \in \mathcal{C}$, $C_1 \cap (\rho(C_1) \cup \rho(C_2)) \neq C_2 \cap (\rho(C_1) \cup \rho(C_2))$ if $C_1 \neq C_2$.



Let $C_i = B_1(v_i)$ and $\mathcal{C} = \{C_1, C_3, C_4\}$.

Let $\rho(C_1) = v_1$,

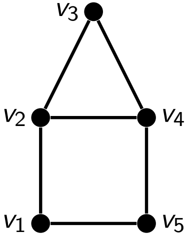
$\rho(C_3) = v_5$,

$\rho(C_4) = v_2$.

Representation Maps

Representation map of size k for a concept class $\mathcal{C}$

A function ρ assigning to each $C \in \mathcal{C}$, a sample X_C of size at most k that is realizable by C, and such that, for any two $C_1, C_2 \in \mathcal{C}$, $C_1 \cap (\rho(C_1) \cup \rho(C_2)) \neq C_2 \cap (\rho(C_1) \cup \rho(C_2))$ if $C_1 \neq C_2$.



Let $C_i = B_1(v_i)$ and $\mathcal{C} = \{C_1, C_3, C_4\}$.

Let $\rho(C_1) = v_1$,

$\rho(C_3) = v_5$,

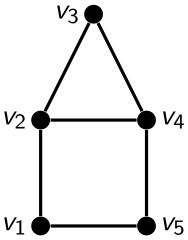
$\rho(C_4) = v_2$.

$$C_1 \cap (\rho(C_1) \cup \rho(C_3)) = \{v_1, v_5\} \neq \emptyset = C_3 \cap (\rho(C_1) \cup \rho(C_3)),$$

Representation Maps

Representation map of size k for a concept class $\mathcal{C}$

A function ρ assigning to each $C \in \mathcal{C}$, a sample X_C of size at most k that is realizable by C, and such that, for any two $C_1, C_2 \in \mathcal{C}$, $C_1 \cap (\rho(C_1) \cup \rho(C_2)) \neq C_2 \cap (\rho(C_1) \cup \rho(C_2))$ if $C_1 \neq C_2$.



Let $C_i = B_1(v_i)$ and $\mathcal{C} = \{C_1, C_3, C_4\}$.

Let $\rho(C_1) = v_1$,

$\rho(C_3) = v_5$,

$\rho(C_4) = v_2$.

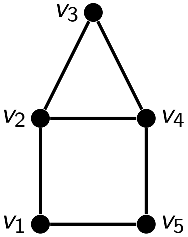
$$C_1 \cap (\rho(C_1) \cup \rho(C_3)) = \{v_1, v_5\} \neq \emptyset = C_3 \cap (\rho(C_1) \cup \rho(C_3)),$$

$$C_1 \cap (\rho(C_1) \cup \rho(C_4)) = \{v_1, v_2\} \neq \{v_2\} = C_4 \cap (\rho(C_1) \cup \rho(C_4)),$$

Representation Maps

Representation map of size k for a concept class $\mathcal{C}$

A function ρ assigning to each $C \in \mathcal{C}$, a sample X_C of size at most k that is realizable by C, and such that, for any two $C_1, C_2 \in \mathcal{C}$, $C_1 \cap (\rho(C_1) \cup \rho(C_2)) \neq C_2 \cap (\rho(C_1) \cup \rho(C_2))$ if $C_1 \neq C_2$.



Let $C_i = B_1(v_i)$ and $\mathcal{C} = \{C_1, C_3, C_4\}$.

Let $\rho(C_1) = v_1$,

$\rho(C_3) = v_5$,

$\rho(C_4) = v_2$.

$$C_1 \cap (\rho(C_1) \cup \rho(C_3)) = \{v_1, v_5\} \neq \emptyset = C_3 \cap (\rho(C_1) \cup \rho(C_3)),$$

$$C_1 \cap (\rho(C_1) \cup \rho(C_4)) = \{v_1, v_2\} \neq \{v_2\} = C_4 \cap (\rho(C_1) \cup \rho(C_4)),$$

$$C_3 \cap (\rho(C_3) \cup \rho(C_4)) = \{v_2\} \neq \{v_2, v_5\} = C_4 \cap (\rho(C_3) \cup \rho(C_4)).$$

Related Work for Representation Maps and Results Included in Part II

Representation maps were defined to characterize and construct USCS for maximum classes [Kuzmin and Warmuth, 2007].

Characterization was extended to ample classes [Chalopin et al., 2019].

Related Work for Representation Maps and Results Included in Part II

Representation maps were defined to characterize and construct USCS for maximum classes [Kuzmin and Warmuth, 2007].

Characterization was extended to ample classes [Chalopin et al., 2019].

We show:

- Representation maps of size 2 for $\mathcal{B}(T)$ for trees T;
- Representation maps of size 2 for $\mathcal{B}(G)$ for interval graphs G;

Rep. maps of size 2 for trees and interval graphs

Trees T:

For any vertex $x \in V(T)$ and any $r \geq 0$, $\rho(B_r(x)) = \{u_x^+, v_x^+\}$, where u_x^+, v_x^+ is a diametral pair of $B_r(x)$.

Interval graphs G:

For any vertex $x \in V(G)$ and any $r \geq 0$, $\rho(B_r(x)) = \{u_x, v_x\}$, where u_x, v_x is a farthest pair of $B_r(x)$.

Further Work

- We are studying SCS and Rep. maps for balls in graphs for cycles and cacti.
- Also, balls of radius 1 for SCS for planar graphs.
- Other graph classes would be interesting.
- **Does every family of balls of VC-dimension d admit SCS of size $O(d)$?**

Graph Class	**Rep. map**	**USCS**	**LSCS**
Trees ($\mathcal{B}(T)$)	≤ 2	≤ 2	
Trees ($\mathcal{B}_r(T)$)	≤ 2		(≤ 2 w/ info) ≤ 6
Interval graphs ($\mathcal{B}(G)$)	≤ 2	≤ 3	
Interval graphs ($\mathcal{B}_r(G)$)	≤ 2		≤ 3
Cycles ($\mathcal{B}(G)$)	≤ 3	≤ 3	
Cacti ($\mathcal{B}(G)$)	≤ 4	?	?

Thanks!