PHOEG Helps Obtaining Extremal Graphs

Gauvain Devillez

Joint work with Hadrien Mélot and Pierre Hauweele

Service d'Algorithmique, Département d'Informatique Faculté des Sciences, Université de Mons

<u>U</u>MONS

ECCO XXIX, May 26, 2016

Introduction

We consider simple undirected graphs.



For a graph G = (V, E),

• its order |V| is denoted by n;

• its size |E| is denoted by *m*.

Introduction

A graph invariant is a function on graphs that is constant on isomorphism classes.

Examples: order *n*, size *m*, chromatic number χ , maximum degree Δ , diameter *D*, planarity, ...

Introduction

A graph invariant is a function on graphs that is constant on isomorphism classes.

Examples: order *n*, size *m*, chromatic number χ , maximum degree Δ , diameter *D*, planarity, . . .

Example (Several isomorphic graphs \rightarrow one graph *G*)



 $n(G) = 4, m(G) = 5, \chi(G) = 3,$

 $\Delta(G) = 3$, D(G) = 2, planarity(G) = true, ...

Extremal Graph Theory

Extremal Graph Theory aims to find bounds on a graph invariant under some constraints.

Generally, those constraints are of two types:

- restricting class of graphs (e.g., connected graphs, trees);
- fixing (and restricting) values of other invariants (e.g., size, maximum degree).

Results in Extremal Graph Theory mainly consists in

- giving bounds;
- characterizing graphs achieving these bounds.

Context: Computer-assisted discovery in Extremal Graph Theory

- Context: Computer-assisted discovery in Extremal Graph Theory
- Several existing systems: Graph, Graffiti, AutoGraphiX, GraPHedron, ...
 - exploit different ideas to help graph theorists

- Context: Computer-assisted discovery in Extremal Graph Theory
- Several existing systems: Graph, Graffiti, AutoGraphiX, GraPHedron, ...
 - exploit different ideas to help graph theorists
- Objectives of this talk:
 - presentation of PHOEG, a successor of GraPHedron
 - use of an illustrative problem (eccentric connectivity index, ECI)

- Context: Computer-assisted discovery in Extremal Graph Theory
- Several existing systems: Graph, Graffiti, AutoGraphiX, GraPHedron, ...
 - exploit different ideas to help graph theorists
- Objectives of this talk:
 - presentation of PHOEG, a successor of GraPHedron
 - use of an illustrative problem (eccentric connectivity index, ECI)
- Remark: work under progress
 - PHOEG is currently a prototype
 - the problem about ECI is not fully solved

Let v be a vertex of a graph G, recall that:

• degree d(v) = number of adjacent vertices of v;



Let v be a vertex of a graph G, recall that:

- degree d(v) = number of adjacent vertices of v;
- eccentricity $\epsilon(v)$ = maximal distance between v and any other vertex.

Example



Definition

The Eccentric Connectivity Index (ECI) of a graph G, denoted by $\xi^{c}(G)$, is

$$\xi^{c}(G) = \sum_{v \in V} d(v)\epsilon(v).$$

Definition

The Eccentric Connectivity Index (ECI) of a graph G, denoted by $\xi^{c}(G)$, is

$$\xi^{c}(G) = \sum_{v \in V} d(v)\epsilon(v).$$



Now, let's make extremal graph theory about ξ^c with the help of a computer.

First step: define a problem by choosing constraints.

Now, let's make extremal graph theory about ξ^c with the help of a computer.

First step: define a problem by choosing constraints.

Several papers containing bounds on ξ^c — using various invariants as constraints — have been published (since 2010). However, the two simplest graph invariants are the order *n* and the size *m* and this leads to the following natural question.

Now, let's make extremal graph theory about ξ^c with the help of a computer.

First step: define a problem by choosing constraints.

Several papers containing bounds on ξ^c — using various invariants as constraints — have been published (since 2010). However, the two simplest graph invariants are the order *n* and the size *m* and this leads to the following natural question.

Problem

Among connected graphs of order *n* and size *m*, what is the maximum possible value for ξ^c ?

(To avoid infinite eccentricities, we restrict the problem to connected graph)

We define $E_{n,m}$ as follows :

n = 7, m = 14

We define $E_{n,m}$ as follows :

 The biggest possible clique without disconnecting the graph.

$$n = 7, m = 14$$



We define $E_{n,m}$ as follows :

- The biggest possible clique without disconnecting the graph.
- A path with the remaining vertices.

$$n = 7, m = 14$$



We define $E_{n,m}$ as follows :

- The biggest possible clique without disconnecting the graph.
- A path with the remaining vertices.
- Add remaining edges between vertices of the clique and the first vertex of the path.

$$n = 7, m = 14$$



We define $E_{n,m}$ as follows :

- The biggest possible clique without disconnecting the graph.
- A path with the remaining vertices.
- Add remaining edges between vertices of the clique and the first vertex of the path.





This graph is unique for given n and m. We define $d_{n,m}$ as the diameter of $E_{n,m}$.



Conjecture of Zhang, Liu and Zhou

Conjecture (Zhang, Liu and Zhou, 2014)

Let G be a graph of order n and size m such that $d \ge 3$. Then,

 $\xi^{c}(G) \leq \xi^{c}(E_{n,m}),$

with equality if and only if $G \simeq E_{n,m}$.

- The authors prove that the conjecture is true when m = n 1, n, ..., n + 4 (if *n* is large enough).
- It exists a "proof" published in a journal of University of Isfahan (Iran, 2014) but that is obviously wrong.

Conjecture of Zhang, Liu and Zhou

Conjecture (Zhang, Liu and Zhou, 2014)

Let G be a graph of order n and size m such that $d \ge 3$. Then, $\xi^{c}(G) \le \xi^{c}(E_{n,m})$, with equality iff $G \simeq E_{n,m}$.

This conjecture leads to several questions:

- Is the conjecture true?
- If yes, how to prove it?
- If no, how to improve or correct it?
- What about graphs such that d < 3?

In the following, we will show how PHOEG can help to study all of the above questions and to raise new ones.

GraPHedron's main principle:

- view graphs as points in the space of invariants;
- compute the convex hull of these points (for small values of *n*).

PHOEG is intended to be the successor of GraPHedron. It can be used to explore graphs' convex hull but also go further (see later).



Е



Е

Polytope for n = 7





Е



Е



Ε

Exploring ξ^c with PHOEG – observations



There is a lot of possible observations on these polytopes.

Observations and questions



How to explain the grid?

■ Is the conjecture of Zhang, Liu and Zhou true when d ≥ 3?

• Upper bound when d < 3?

Observations and questions



- How to explain the grid? GraPHedron: gives no access to inner points
- Is the conjecture of Zhang, Liu and Zhou true when d ≥ 3? GraPHedron does not allow

to constraint points

■ Upper bound when *d* < 3? Idem

These questions are outside the scope of the former system: let's dive into PHOEG!

From GraPHedron to PHOEG

- Former system: graphs and invariant's values written sequentially in files;
- PHOEG uses a PostgreSQL DB with more than 12 million of non-isomorphic graphs (up to order 10);
- Each graph has its unique signature:
 - to each graph *G*, one assigns a representative of its isomorphism class;
 - it is called the canonical form of *G*;
 - in practice, Canon(G) is the smallest graph in the isomorphism class of G (in the lexicographical order induced by adjacency matrices);
 - the canonical matrix is then translated into a string (graph6 format):

$$sig(C_5) = "DqK";$$

$$sig(K_3) = "Bw".$$

This allows complex (and fast) queries on graphs.
Invariants' Database

Graphs		NumVertices]	NumEdges			ECI	
signature	1	signature	val	1	signature	val	1	signature	val
A_		A_	2]	A_	1]	A_	2
Α?		A?	2		A?	0		BW	6
B?		B?	3		B?	0		Bw	6
BG		BG	3		BG	1		C^	14
Bw		Bw	3		Bw	3		C~	12
BW		BW	3		BW	2		CF	9
Cʻ		Cʻ	4		Cʻ	2		CN	13
C^		C^	4		C^	5		Cr	16
C~		C~	4		C~	6		CR	14
C?		C?	4		C?	0		D'[25
C@		C@	4		C@	1		D'{	20

Database query – Points and multiplicities

```
SELECT P.val AS eci, num_edges.val AS m,
                                              eci l m
                                                        | mult
  COUNT(*) AS mult
                                               ____+
FROM eci P
                                                47 I
                                                      8 |
                                                             5
                                                46 I
                                                      8 |
                                                             3
  JOIN num_vertices USING(signature)
  JOIN num_edges USING(signature)
                                                             3
                                                40 I
                                                      8 |
WHERE num_vertices.val = 7
                                                32 |
                                                      7 |
                                                             3
GROUP BY m, eci;
                                                48 | 12 |
                                                            55
                                                48 I
                                                     18 I
                                                             1
                                                61 I
                                                     14 I
                                                             4
                                                59 | 13 |
                                                             1
                                                48 | 11 |
                                                            17
                                                43 | 9 |
                                                            14
                                                47 I
                                                      6 |
                                                             1
                                                64 | 10 |
                                                             1
                                                59 I
                                                     11 |
                                                             1
                                                45 | 9 |
                                                             7
                                                             2
                                                38 | 6 |
                                                    [...]
```

Database query – Polytope

SELECT ST_AsText(ST_ConvexHull(ST_Collect(ST_Point(eci, m))))
FROM poly;

st_astext POLYGON((18 6,42 21,66 18,68 17,66 11,62 8,54 6,18 6)) Database query – Polytope



Ε

Database query – adding other information

SELECT num_edges.val AS m, p.val AS eci, d.val AS d, diam.val AS diam FROM eci p JOIN num_vertices USING(signature) JOIN num_edges USING(signature) JOIN d USING(signature) JOIN diam USING(signature) WHERE num_vertices.val = 7 ORDER BY diam, d, m, eci;

m e		eci	I	d		diam	
	+-		+-		+-		
21	Ι	42	T	1	Ι	1	
16	T	46	Т	2	Τ	2	
16	T	52	I	2	Ι	2	
16	T	52	I	2	Ι	2	
16	Т	52	I	2	Ι	2	
16	Т	52	I	2	Ι	2	
16	Т	52	I	2	Ι	2	
16	T	58	I	2	Ι	2	
16	Т	58	I	2	Ι	2	
16	Т	58	I	2	Ι	2	
16	Т	58	I	2	Ι	2	
16	T	58	I	2	Ι	2	
16	T	58	I	2	Ι	2	
16	Т	58	I	2	Ι	2	
16	T	58	I	2	Ι	2	
		[.		.]			

Coloring points with values of d



Recall that the conjecture is stated for $d \ge 3$. Is it true for n = 7?

G. Devillez - UMONS

PHOEG Helps Obtaining Extremal Graphs

ECCO XXIX - 2016 24 / 38

Database query – Extremal graphs

```
WITH tmp AS (
  SELECT n.val AS n, m.val AS m,
    P.signature, P.val AS eci, d.val AS d
   rank() OVER (
      PARTITION BY n.val, m.val
      ORDER BY P.val DESC
    ) AS pos
 FROM num vertices n
  JOIN num_edges m USING(signature)
  JOIN d USING(signature)
  JOIN eci P USING(signature)
  WHERE n_val = 7
)
SELECT signature AS sig, n, m, eci, d
FROM tmp
WHERE pos = 1 AND d \geq 3
ORDER BY n, m, d, eci;
```

| n | m | eci | d sig ---+---+------F@IQO | 7 | 6 I 54 I 6 F@ʻJ 7 | 7 | 57 5 FgCXW | 7 | 8 | 62 5 FWCYw | 7 | 9 | 62 Т 4 FgCxw | 7 | 10 | 64 4 F'Kvw | 7 | 11 | 66 4 $F'K_{ZW} \mid 7 \mid 12 \mid$ 65 3 F'Lzw | 7 | 13 | 65 Ι З F'\zw | 7 | 14 | 65 З F.I] | w | 7 | 15 | 65 I 3 FJ\|w | 7 | 15 | 65 I 3

Database query – Extremal graphs

```
WITH tmp AS (
                                             sig | n | m | eci | d
  SELECT n.val AS n, m.val AS m,
                                                   ___+
    P.signature, P.val AS eci, d.val AS d
                                            F@IQO | 7 |
    rank() OVER (
                                            F@'J_ | 7 | 7 |
     PARTITION BY n.val, m.val
                                            FgCXW | 7 | 8 |
     ORDER BY P.val DESC
                                            FWCYw | 7 | 9 |
    ) AS pos
                                            FgCxw | 7 | 10 |
                                            F'Kvw | 7 | 11 |
 FROM num vertices n
  JOIN num_edges m USING(signature)
                                            F'Kzw | 7 | 12 |
  JOIN d USING(signature)
                                            F'Lzw | 7 | 13 |
  JOIN eci P USING(signature)
                                            F'\zw | 7 | 14 |
  WHERE n_val = 7
                                           F.I] | w | 7 | 15 |
)
                                            FJ\|w | 7 | 15 |
SELECT signature AS sig, n, m, eci, d
FROM tmp
WHERE pos = 1 AND d \geq 3
ORDER BY n, m, d, eci;
             \Rightarrow counter-example to the conjecture !
             Extremal graphs are not always unique
```

G. Devillez - UMONS

PHOEG Helps Obtaining Extremal Graphs

6 I

54 I 6

57 I 5

62 4

64 L 4

66

65

65 З

65

65 I 3

65 | 3

62

5

4

3

З

Counter-example (n = 7 and m = 15)



Counter-example (n = 7 and m = 15)



Counter-example (n = 7 and m = 15)



It is possible to construct counter-examples for any values of $n \ge 6$ (with d = 3).

Coloring points with values of d



Upper bound when $d \leq 2$?

PHOEG Helps Obtaining Extremal Graphs

ECCO XXIX - 2016 27 / 38

Upper facet of the polytope (n = 7)



A new upper bound tight when $d \leq 2$

Theorem

Let G be a graph of order n and size m. Then,

$$\xi^{c}(G) \leq n(n-1)(n-2) - 2m(n-3),$$

with equality if and only if G is the complement of a matching.

Note that the bound is valid for all graphs but can be tight only if

$$m \ge \binom{n}{2} - \left\lfloor \frac{n}{2} \right\rfloor,$$

(and thus $d \leq 2$).

Coloring points with values of the diameter



Coloring points with values of the diameter



Can the diameter explain the blue grid? Actually, yes!

G. Devillez - UMONS

PHOEG Helps Obtaining Extremal Graphs

ECCO XXIX - 2016 30 / 38



Suppose D(G) = 2 (light blue points)
For each vertex v, since D(G) = 2, either e(v) = 1 or e(v) = 2



- Suppose D(G) = 2 (light blue points)
- For each vertex v, since D(G) = 2, either ϵ(v) = 1 or ϵ(v) = 2
- If $\epsilon(v) = 1$, then v is dominant and d(v) = n 1



- Suppose D(G) = 2 (light blue points)
- For each vertex v, since D(G) = 2, either ϵ(v) = 1 or ϵ(v) = 2
- If $\epsilon(v) = 1$, then v is dominant and d(v) = n 1
- Let k be the number of dominant vertices of G



- Suppose D(G) = 2 (light blue points)
- For each vertex v, since D(G) = 2, either ϵ(v) = 1 or ϵ(v) = 2
- If $\epsilon(v) = 1$, then v is dominant and d(v) = n 1
- Let k be the number of dominant vertices of G
 The sum of degrees of non dominant vertices is

$$2m-k(n-1)$$



- Suppose D(G) = 2 (light blue points)
- For each vertex v, since *D*(*G*) = 2, either $\epsilon(v) = 1$ or $\epsilon(v) = 2$
- If $\epsilon(v) = 1$, then v is dominant and d(v) = n 1
- Let k be the number of dominant vertices of G
 The sum of degrees of non dominant vertices is
 2m - k(n - 1)

Thus,

$$\xi^{c}(G) = k(n-1) + 2(2m - k(n-1)) = 4m - k(n-1),$$

that is maximum if k = 0 and, moreover, explain the grid.

PHOEG – Transproof

Up to this point, we have

- a tight upper bound when $d \leq 2$;
- and counter-examples for the unicity if d = 3.

However, the conjecture may be true if $d \ge 4$ (actually, we believe it is).

Is PHOEG able to help also for a proof?

PHOEG – Transproof

Up to this point, we have

- a tight upper bound when $d \leq 2$;
- and counter-examples for the unicity if d = 3.

However, the conjecture may be true if $d \ge 4$ (actually, we believe it is).

Is PHOEG able to help also for a proof?

This is the purpose of the *Transproof* module:

- using graph transformations is a common proof technique;
- not always easy to find "good" transformations.

























Graph database



Metagraph stored in a graph DB (Neo4j)

```
Easy queries, e.g.,
match (n)-[e:EdgeRemoval]->(m)
where n.invariant < m.invariant
return n,e,m
```
Finding transformations

Actually, removing an edge is not well suited for our problem.

Indeed, the size of our graphs is fixed.

Actually, removing an edge is not well suited for our problem.

Indeed, the size of our graphs is fixed.

A rotation can be better since it keeps the size unchanged.

Definition

Let G = (V, E) be a graph and u, v, w be three vertices of G such that $uv \in V$ and $uw \notin V$. Then, G' is the graph obtained from G by applying a *rotation* rot(u, v, w) if

$$G'=G-uv+uw.$$

и

Actually, removing an edge is not well suited for our problem.

Indeed, the size of our graphs is fixed.

A rotation can be better since it keeps the size unchanged.

Definition

Let G = (V, E) be a graph and u, v, w be three vertices of G such that $uv \in V$ and $uw \notin V$. Then, G' is the graph obtained from G by applying a *rotation* rot(u, v, w) if

$$G' = G - uv + uw.$$

w

The metagraph of rotations for ξ^c when n = 5 et m = 6



The metagraph of rotations for ξ^c when n = 5 et m = 6



Applying only one rotation is thus not sufficient to have a proof.

Finding good transformations for ξ^c : work in progress.

- Not only extremal graphs are useful to study extremal properties of an invariant
- Exact approach limited to small graphs ($n \le 10$)
- However, dealing with small graphs has already shown to be very useful and led to numerous results (AutoGraphiX, GraPHedron)

Perspectives

- Invariants' DB allows a form of dynamic programming
- Create a simple interface for queries
- Allow easy visualization and manipulation of outputs (GUI, PDF, etc.)
- Simplify the definitions of transformations
- Suggest automatically (a short list of) transformations

Appendix

Eccentric Connectivity Index

History and motivation

- Sharma, Goswani and Madan introduced ξ^c in 1997 in Chemistry;
- Useful as a discriminating topological descriptor for Structure Properties and Structure Activity studies;
- Since 1997, more than 200 chemical papers about ξ^c: applications in drug design, prediction of anti-HIV activities, etc.

Eccentric Connectivity Index

History and motivation

- Sharma, Goswani and Madan introduced ξ^c in 1997 in Chemistry;
- Useful as a discriminating topological descriptor for Structure Properties and Structure Activity studies;
- Since 1997, more than 200 chemical papers about ξ^c: applications in drug design, prediction of anti-HIV activities, etc.
- However, the first mathematical paper with extremal properties on ξ^c was published only in 2010;
- Since 2010, about a dozen papers containing bounds on ξ^c.

Definition

For positive integers *n* and *m* with $n-1 \le m \le \binom{n}{2}$, let

$$d_{n,m} = \left\lfloor \frac{2n+1-\sqrt{17+8(m-n)}}{2} \right\rfloor$$

In the following, we simply use d for $d_{n,m}$.

Definition

Let $E_{n,m}$ be the graph obtained from a clique K_{n-d-1} and a path $P_{d+1} = v_0 v_1 \dots v_d$ by joining each vertex of the clique to both v_d and v_{d-1} , and by joining

$$m-n+1-\binom{n-d}{2}$$

vertices of the clique to v_{d-2} .

т	4	5	6	7	8	9	10
d	4	3	3	2	2	2	1
n-d-1	0	1	1	2	2	2	3
$\#$ edges to v_{d-2}	0	0	1	0	1	2	0







т	4	5	6	7	8	9	10	
d	4	3	3	2	2	2	1	
n-d-1	0	1	1	2	2	2	3	
$\#$ edges to v_{d-2}	0	0	1	0	1	2	0	



т	4	5	6	7	8	9	10
d	4	3	3	2	2	2	1
n-d-1	0	1	1	2	2	2	3
$\#$ edges to v_{d-2}	0	0	1	0	1	2	0

т	4	5	6	7	8	9	10
d	4	3	3	2	2	2	1
n-d-1	0	1	1	2	2	2	3
$\#$ edges to v_{d-2}	0	0	1	0	1	2	0

