Antimatroids and Convex geometries
MPRI 2017–2018

22 décembre 2017

Schedule

Algorithm design is hard

1. No general results
2. It seems that there are only ad hoc solution
3. If we have a little change in the problem the solution can be drastically changed (from P to NP-complete for example)

Some answers

1. Still not enough general theorems for algorithm design
2. The two last points are also true for complexity theory.
3. Maximum sized maximal matching (P) versus Minimum size maximal matching (NP-hard).
4. Modules (or common intervals) with a fixed number of errors.

One of the ideas to navigate in this complex world of algorithms

Study the underlying algebraic structure

Matroidal Example : Forest as edge-sets in a graph, vector spaces.
Exchange property between basis. Greedy algorithms

Matroidal again Matchings = intersection of two matroids
Family close by intersection and exchange property along the symmetric difference
Algorithms using a series of improvements by interchanges (matchings, maximum flows . . .)

Partitive and variants Many combinatorial decompositions (modular decomposition, split decomposition . . .)

Study the underlying algebraic structure

Dynamic programming There exists a polynomial recursive scheme to solve the problem

Some boarder line

Moore families Independent sets in a graph.
Just closed under intersection. NP-hard

Moore families Dominating sets in a graph
Just closed under union NP-hard

In the two last items, only recursive schemes in $O(a^n)$ with $a > 1$ are known.

Considering a new algorithmic problem, there is no way to avoid :

1. To try to transform the problem into some known one
2. Try on many examples, using a pencil or a computer in order to obtain some feelings on the problem.

The objective is to understand algorithmic greediness, and in particular why it works so well in applications.

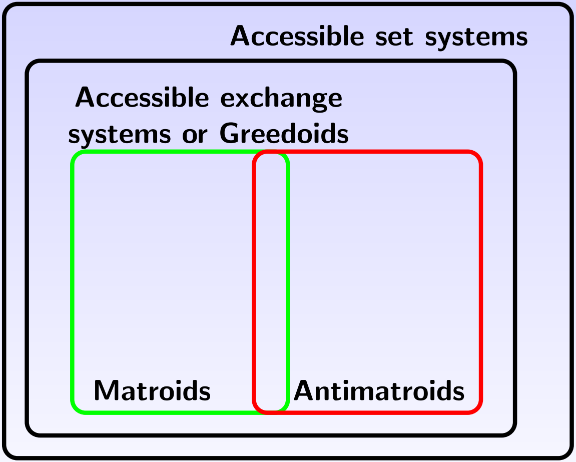


FIGURE: Greediness Landscape

En français le Glouton ou Carcajou

Accessible exchange systems also called greedoids

$GR = (V, \mathcal{F})$ is an accessible exchange system or greedoid :
a finite set V, $\mathcal{F}$ a family of subsets of V that satisfies :

1. $\emptyset \in \mathcal{F}$
2. If $X \in \mathcal{F}$ such that $X \neq \emptyset$ then $\exists x \in X$ such that $X \setminus x \in \mathcal{F}$
3. If $X, Y \in \mathcal{F}$ and $|X| > |Y|$ then $\exists x \in X \setminus Y$ such that $Y \cup \{x\} \in \mathcal{F}$.

These families are also called greedoids
they generalizes matroids.

As a consequence :
All maximal (under inclusion) elements of $\mathcal{F}$ have the same size are called basis. This common size is called the rank of the greedoid.

A matroid by its independents

Withney [1935] : Let $M = (X, \mathcal{I})$ be a set family where X is a finite ground set and $\mathcal{I}$ a family of subsets of X.
M is a matroid if it satisfies the 3 following axioms :

1. $\emptyset \in \mathcal{I}$
2. If $I \in \mathcal{I}$ alors $\forall J \subseteq I$, $J \in \mathcal{I}$
3. If $I, J \in \mathcal{I}$ et $|I| = |J| + 1$ then $\exists x \in I$-J such that : $J \cup \{x\} \in \mathcal{I}$.

Variations around axiom 3

- 3' If $I, J \in \mathcal{I}$ and $|I| > |J|$ then $\exists x \in I$-J such that : $J \cup \{x\} \in \mathcal{I}$.
- 3" $\forall A \subseteq X$ and $\forall I, J \subseteq A$, $I, J \in \mathcal{I}$ and maximal then $|I| = |J|$
- It is easy to be convinced that 3, 3' and 3" are equivalent.

To be an hereditary family is stronger than axiom 2 of greedoids. Thus every matroid is a greedoid. Furthermore matroids are exactly the hereditary greedoids.

Examples of greedoid :

Rooted trees systems or undirected branchings

For an undirected connected graph G and a vertex $r \in G$, we define the set systems :
$(E(G), \mathcal{F})$
Where $F \in \mathcal{F}$ iff F is the edge set of a tree containing r.

Rooted trees gredoids

Axiom 1 of greedoids is trivially satisfied, because r itself is a tree rooted in r.
If F defines a tree and has at least one edge, then it contains least two leaves x, y. Say $x \neq r$ and e_x its attached edge.
$F \setminus e_x$ is a tree (eventually empty is $y = r$). Therefore axiom 2 is satisfied.

Axiom 3

Let X, Y be two subtrees rooted in r such that : $|X| > |Y|$.
If we denote by $span(X)$, resp. $span(Y)$ the set of vertices adjacent to at least one edge of X, resp. Y.
$|span(X)| > |span(Y)|$ and therefore if we consider the cut C yielded by $span(Y)$ in $span(X)$ around the component containing r, there exists at least one edge e in this cut and $Y + e$ is a tree containing r.
So $Y + e \in \mathcal{F}$.

Theorem

Undirected rooted branchings are greedoids.
A basis of an undirected branching corresponds to the set of edges of a graph search in G starting at r.

Rooted trees are not matroids

Although forests in a graph yield a matroid, rooted trees does not share the hereditary property of matroids, since deleting an edge could disconnect the rooted tree.

Rooted branchings in directed graphs

Rooted branchings

For a directed connected graph G and a vertex $x_0 \in G$, we define the set systems :
$(A(G), \mathcal{F})$
Where $F \in \mathcal{F}$ iff $F \subseteq A(G)$ is the arc set of an arborescence containing x_0 or equivalently for each vertex $x \in F$ there is a unique path from x_0 to x.

Rooted trees gredoids

Axiom 1 of greedoids is trivially satisfied, because r itself is a tree rooted in r.
If F is an arborescence and has at least one arc, then it contains least two leaves x, y. Say $x \neq x_0$ and a_x its attached arc.
$F \setminus a_x$ is an arborescence (eventually empty is $y = x_0$). Therefore axiom 2 is satisfied.

Axiom 3

Let X, Y be two arborescences rooted in x_0 such that : $|X| > |Y|$.
If we denote by $span(X)$, resp. $span(Y)$ the set of vertices adjacent to at least one arc of X, resp. Y.
$|span(X)| > |span(Y)|$ and therefore if we consider the cut C yielded by $span(Y)$ in $span(X)$ around the component containing x_0, there exists at least one arc a in this cut and $Y + a$ is a tree containing x_0.
So $Y + a \in \mathcal{F}$.

Theorem

Directed rooted branchings are greedoids.
$Br = (V(G), \mathcal{F})$ is a directed branching greedoid if
$\mathcal{F}$ is the set of all branchings rooted in x_0 of G.
A basis of an undirected branching corresponds to the set of arcs of a graph search in G starting at x_0.

Directed branchings are not matroids

Since they does not share the hereditary property of matroids, since deleting an edge could disconnect the branching.

Antimatroids another example of greedoids

$AM = (V, \mathcal{F})$ is a antimatroid :
a finite set V, $\mathcal{F}$ a family of subsets of V that satisfies :

1. $\emptyset \in \mathcal{F}$
2. If $X \in \mathcal{F}$ such that $X \neq \emptyset$ then $\exists x \in X$ such that $X \setminus x \in \mathcal{F}$
3. If $X, Y \in \mathcal{F}$ and $X \not\subseteq Y$ then $\exists x \in X \setminus Y$ such that $Y \cup \{x\} \in \mathcal{F}$.

Remarks :

- Axiom 3 implies $X \cup Y \in \mathcal{F}$.
- Antimatroids yield another particular case of greedoids
- Using axiom 3, if $\mathcal{F}$ is finite then it admits a unique maximal element (the union of all elements in $\mathcal{F}$).

Example

Let G be a chordal graph.
$\mathcal{F}= \{X \subseteq V(G) | X$ is the beginning of a simplicial elimination scheme of $G\}$
$AM = (V, \mathcal{F})$ is an antimatroid.

proof

Axiomes 1 and 2 are trivially satisfied by $\mathcal{F}$.
Let us consider the third one. First we use the Y as the starting of a simplicial elimination of G. Consider the ordering τ of X which is the starting of a simplicial elimination scheme of G. Let x_1 the first element of $X \setminus Y$ with respect to τ.
Necessarily after Y we can use x_1 as a simplicial vertex in the remaining graph.
Therefore $Y \cup \{x\}$ is the beginning of a simplicial elimination scheme of G.

Similarly Ideals (resp. filters) of a partial order yield an antimatroid.

Optimisation on simplicial elimination schemes

For a chordal graph G :
Let $\tau = x_1, \ldots x_n$ be a simplicial elimination scheme.
$Bump(\tau) = \#$ consecutive pair of elements in τ which are not adjacent in the graph G.
$Bump(G) = min_\tau Bump(\tau)$
Can we compute in a greedy way a simplicial elimination scheme with a minimum number of bump ?
Seems to be related to the leafage problem :
We have $Bump(G) = 0$ if G is an interval graph and more generally :
$Bump(G) \leq Leafage(G) - 2$.

Accessible systems

$(V, \mathcal{F})$ is an accessible system if it satisfies the following condition : a finite set V, $\mathcal{F}$ a family of subsets of V that satisfies :

1. $\emptyset \in \mathcal{F}$
2. If $X \in \mathcal{F}$ such that $X \neq \emptyset$ then $\exists x \in X$ such that $X \setminus x \in \mathcal{F}$

A personal question

How to evaluate the proportion of accessible systems or quasi accessible systems among all set families ?

This could explain, why greedy algorithms work so well in practice.

Convex spaces from Kay and Wamble 1971

$(V, \mathcal{F})$ where V is finite set and $\mathcal{F}$ a family of subsets of V, is a convex space if it satisfies the following conditions :

1. $\emptyset \in \mathcal{F}$, $V \in \mathcal{F}$.
2. $\mathcal{F}$ is closed under intersection

Convex hull

We define for $\forall X \in V$, the convex hull of X denoted by $\tau_{\mathcal{F}}(X)$ as the intersection of all convex supersets of X

Also called Closure systems

For a ground set V, the application $\mathcal{K} \subseteq \in^{\mathcal{V}}$ is a closure system or Moore family if it satisfies :

1. $V \in \mathcal{K}$
2. $\forall X, Y \in \mathcal{K}$ implies $X \cap Y \in \mathcal{K}$.

i.e. a family closed under intersection.

Closure operator

For a ground set V, the application $\sigma : 2^V \to 2^V$ is a closure operator if it satisfies :

1. $\sigma(\emptyset) = \emptyset$
2. $\forall X \subseteq V$, $X \subseteq \sigma(X)$.
3. $\forall X, Y \subseteq V$ with $X \subseteq Y$ then $\sigma(X) \subseteq \sigma(Y)$.
4. $\forall X \subseteq V$, $\sigma(\sigma(X)) = \sigma(X)$.

Example :

The convex application $\tau_{\mathcal{F}}$ is a closure operator.

Equivalence :

There is a one-to-one equivalence between the closure systems and the closure operators.

Extreme points for a closure operator

Extreme points

A map $ext : 2^V \rightarrow 2^V$ such that :
$\forall A \subseteq V$, $ext(A) = \{x \in A : x \notin \sigma(A\text{-}x)\}$
is called an extreme point operator of a closure system.
An element in $ext(A)$ is called an extreme element in A.

In other words

If C is a convex set, and x extreme for C, then C-x is still convex.

Convex geometries as defined by Edelman and Jamison 1986

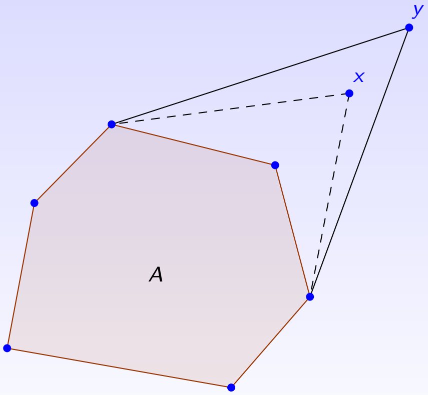
$(V, \mathcal{F})$ where V is finite set and $\mathcal{F}$ a family of subsets of V, is a convex geometry if it satisfies the following conditions :

1. $(V, \mathcal{F})$ is a convex space.
2. The anti-exchange property
   If $X \subseteq V$ and x, y are distincts elements outside of $\tau_{\mathcal{F}}(X)$, then at most one of $x \in \tau_{\mathcal{F}}(X \cup \{y\})$ and $y \in \tau_{\mathcal{F}}(X \cup \{x\})$ holds true.

Closed sets

Let us define closed sets $X \subseteq V$ such that $\tau_{\mathcal{F}}(X) = X$.

Anti-exchange property



Anti-exchange property :
Let $x, y \notin \sigma(A)$ and $x \in \sigma(A + y)$,
then $y \notin \sigma(A + x)$

Complement of antimatroids are convex geometries

$AM = (V, \mathcal{F})$ is an antimatroid :
a finite set V, $\mathcal{F}$ a family of subsets of V that satisfies :

1. $\emptyset \in \mathcal{F}$,
2. If $X \in \mathcal{F}$ such that $X \neq \emptyset$ then $\exists x \in X$ such that $X \setminus x \in \mathcal{F}$,
3. $\mathcal{F}$ is closed under union.

Family of complements

Let $AM = (V, \mathcal{F})$ be an antimatroid, we define :
If $\cup_{F \in \mathcal{F}} F = U$. $CG(AM) = (V, \mathcal{G})$ where $\mathcal{G} = \{U \setminus F | F \in \mathcal{F}\}$

Theorem ??

AM is an antimatroid iff $CG(AM)$ is a convex geometry.

Proof

First $CG(AM)$ is a convex space. Since antimatroids families are closed union their complement are closed under intersection. Therefore $\emptyset$ and U are in $\mathcal{G}$.

End of the proof

The anti-exchange property allows to define for every set $S \subseteq V$ a partial order among x, y distincts elements outside of $\tau_{\mathcal{F}}(S)$.
$x \leq_S y$ if $x \in \tau_{\mathcal{F}}(S + y)$. ($\leq_S$ is clearly transitive).
Therefore for every minimal element x, necessarily : $S + x$ is closed. Else $S + x \subsetneq \tau_{\mathcal{F}}(S + x)$. Therefore there exists $z \in \tau_{\mathcal{F}}(S + x)$-$S + x$ and $z \leq x$, a contradiction.
So for every closed set S, except V there exist x such that $S + x$ is closed. This is dual of accessible property of antimatroids.

Betweenness in usual metric spaces

In a metric space, point y is said to lie between points x and z iff $dist(x, y) + dist(y, z) = dist(x, z)$

Abstract betweenness from V. Chvátal

- Any ternary relation $\mathcal{B}$ such that :
  1. $(A, B, C) \in \mathcal{B}$ implies A, B, C are all distincts,
  2. $(A, B, C) \in \mathcal{B}$ implies $(C, B, A) \in \mathcal{B}$,
  3. $(A, B, C) \in \mathcal{B}$ implies $(C, A, B) \notin \mathcal{B}$.
- Convexity defined by betweenness :
  A set is called convex if, with every two points A and C, it includes all the points B that lie between A and C.
- Not all such convex spaces are convex geometries.

Some convexities on graphs defined with an abstract betweenness relation

For 2 unrelated vertices x, y in G, we define :

- Geodesic convexity (also called Geodetic convexity) $Interval(x, y) = \{z | z \in p \text{ a shortest path from } x \text{ to } y\}$. This yields the geodesic-convexity or $\mathcal{G}$-convexity.
- Monophonic convexity $Interval(x, y) = \{z | z \in p \text{ an induced path from } x \text{ to } y\}$ N.B. since induced paths have no chord, therefore only one sound, so "monophonic" !
- M3 convexity $Interval(x, y) = \{z | z \in p \text{ an induced path of length} \geq 3 \text{ from } x \text{ to } y\}$

- 2-paths-convexity
  $Interval(x, y) = \{z | \exists p, q$ two induced paths, such that p from z to x avoids $N[y]$ and q from z to y avoids $N[x]\}$.
- Induced paths means without chord
- It should be noticed that in the 2-paths-convexity necessarily from the definition : $xy \notin E$ and the paths p, q have length ≥ 2. Therefore (x, y, z) forms an independent triple. But the 2 paths may have a common start from z.

Convexity in partially ordered sets

A set is called convex if, with every two points A and C, it includes all the points B for which $A < B < C$ or $C < B < A$.

Folklore

This interval convexity is a convex geometry on partial orders.
Simple direct proof.

Monophonic convexity

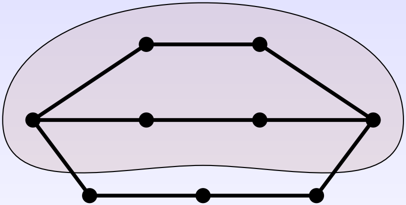
$\mathcal{M}$-convex sets

$C \subseteq V(G)$ is $\mathcal{M}$-convex if every induced path joining two vertices of C is in C.

Observation

A point is extreme for monophonic convexity iff it has two non adjacent neighbours (it is the middle of a P_3).

Geodesic-convexity



A set $C \subseteq V(G)$ is $\mathcal{G}$-convex (geodesic-convex) if every shortest path joining two vertices of C is in C.

Carathéodory numbers. From Carathéodory's theorem 1907

A convex geometry is said to have Carathéodory number k iff every point of every convex set C belongs to the convex hull of some set of at most k extreme points of C.

Folklore

Interval convexity in partially ordered sets is a convex geometry of Carathéodory number 2.

Theorem (Howorka 1977)

Geodesic-convexity is a convex geometry on a graph G iff G is a Ptolemaic graph (i.e., chordal and distance hereditary).

Theorem (Farber and Jamison 1986)

Monophonic convexity yields a convex geometry on a graph G iff G is a chordal graph. This convex geometry as Carathéodory number 2.

M3-convexity

Observation

A point is extreme for M3-convexity iff it is not the middle of a P_4.

Theorem (Farber and Jamison 1986)

If G is HDD-free (House, Domino, Diamond), M3-convexity is a convex geometry.

Theorem (Dragan, Nicolai, Brandstädt 1999)

Monophonic convexity yields a convex geometry on a graph G iff G is weak polarizable (every subgraph is chordal or it has a proper homogeneous set). This convex geometry as Carathéodory number 2.

Theorem

If G is chordal, the last vertex of a LBFS (resp. LDFS) is extremal for the monophonic convexity.

Proof

A simplicial vertex is clearly extreme for monophonic convexity.

Theorem

If G is AT-free, the last vertex of a LBFS is extremal for the 2paths-convexity.

Proof

Direct proof. If the last vertex z belongs to some 2-paths interval $[x, y]$, this would imply an asteroidal triple (x, y, z). Using the unique path in the LBFS tree from x to y. Trivially this path avoids the neighbourhood of z.

LBFS ends at extreme vertices

1. Tarjan et al., G is Chordal, every LBFS ends at a simplicial vertex.
   An extreme vertex for the Monophonic convex geometry on chordal graphs.
2. Dragan et al., G is HDD-free, (House, Domino, Diamond) every LBFS ends at a semi-simplicial vertex.
   Semi-simplicial means not middle vertex of a P_4. An extreme vertex for the M3 convex geometry on HDD-free graphs.
3. Dragan et al., G is interval, every LBFS ends at a diametral simplicial vertex.
4. MH et al., G is cocomp, every LBFS ends at a source of $\overline{G}$.
5. Corneil et al., G is AT-free, every LBFS ends at an extreme vertex for the 2-path convex geometry on AT-free graphs.
6. ...

Perhaps there exists a meta-theorem of the following type

If LBFS "supports" some convexity on an hereditary class of graphs $\mathcal{C}$ then
$\forall G \in \mathcal{C}$, any LBFS applied on G ends at an extreme point of this convexity.

Of course there exist other combinatorial convexities not supported by LBFS.
For strongly chordal graphs, they admit a strong elimination ordering, which can be associated with a convex geometry, but LBFS does not always provide such an ordering.
Such an ordering can be obtained via a lexical ordering.

Research problem

Study the relationships between graph searches and convexity in graphs.

This explain why LBFS is so used as a preprocessing for transitive orientation, modular decomposition, split decomposition, circular-arc recognition…
Notion de certificat

Let us call a shelling for a given convex geometry $\mathcal{H}$, a total order on the elements $x_1, \ldots, x_n$ that satisfies : $\forall i$, x_i is extreme for $\mathcal{H}$ in the subgraph $G\{x_{i+1}, \ldots, x_n\}$.

1. x_i is simplicial in $G\{x_{i+1}, \ldots, x_n\}$ and G is chordal.
2. x_i is simplicial in $T\{x_{i+1}, \ldots, x_n\}$ and T is a tree.
3. x_i is a source (resp a sink) in $P\{x_{i+1}, \ldots, x_n\}$ and P is a partial order.
4. x_i is a false twin (or a true twin) of some $x_j \in G\{x_{i+1}, \ldots, x_n\}$ and G is a cograph.
5. x_i is a pending vertex or is a false twin (or a true twin) of some $x_j \in G\{x_{i+1}, \ldots, x_n\}$ and G is a Distance Hereditary graph
6. ...

Submodular functions

Definition

E a finite ground set, and $f : 2^E \to \mathcal{R}$. f is submodular if for all $\forall A, B \subseteq E$, $f(A \cup B) + f(A \cap B) \leq f(A) + f(B)$

Examples

- Cut functions in graphs and hypergraphs (a case by case proof)
- Number of splitters of a module (a case by case proof)
- Rank functions of matroids
- Any positive linear combination of submodular functions is still submodular.

Minimizing a submodular functions is most of the time polynomial.
Examples (flows , modules, . . .).
Maximizing is more difficult.
Recent uses in Machine Learning Theory.

Polymatroids from Jack Edmonds

A polymatroid is a set systems for which the rank function is a monotone non-decreasing submodular function
verifying :
$f(\emptyset) = 0$
$\forall A, B \subseteq E$, $A \subseteq B$ implies $f(A) \leq f(B)$
Greedy algorithms can be defined in this framework.

Jou-ming Chang, Ton Kloks and Hung-Lung Wang
Gray codes for AT-free Orders via Antimatroids.
IWOCA 2015, LNCS 9538, 77-87, 2016.

Vasek Chvátal
Antimatroids, betweenness, convexity.
Bonn Workshop of Combinatorial Optimization, 57-64, 2008.

Brenda L. Dietrich.
Matroids and antimatroids - a survey.
Discrete Mathematics, 78 : 223-237, 1989.

Feodor Dragan, Falk Nicolai and Andreas Brandstadt,
Convexity and HHD-free graphs.
SIAM J. Discrete Math. vol. 12, 119-135, 1999.

Kenji Kashiwabara and Masataka Nakamura
Antimatroids and convex geometries - The fundamentals.
Dept. of Systems Science, University of Tokyo, Draft 2011.