Representation of simplicial complexes

Dorian Mazauric

Inria Sophia Antipolis - Méditerranée, Geometrica

Motivations and Problems

Simplicial complexes are used extensively in combinatorial and computational topology.

Problem $\rightarrow$ the size of the complexes is very large and the use of simplicial complexes is limited in practice.

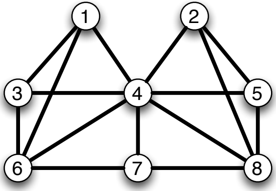
Need for a compact structure $\rightarrow$ an important problem is to store simplicial complexes by using compact structures.

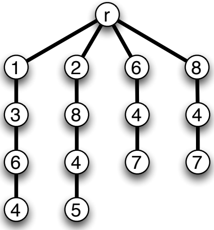
Tree representation $\rightarrow$ every maximal simplex is represented by a path between the root and a leaf.

Definition

Let $\mathcal{K}$ be a simplicial complex and let $\mathcal{E}$ be the set of maximal simplices of $\mathcal{K}$. A tree $T = (V, E, L_1)$ rooted at $r \in V$, $L_1 : V \to \mathcal{V}$, is a tree representation of $\mathcal{K}$ if and only if

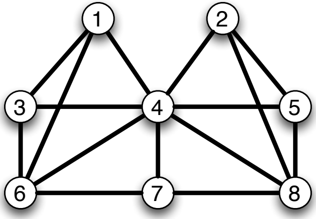
1. $\forall e \in \mathcal{E}$, there is a simple path P_e in T between r and a leaf s. t. $\forall v \in e$, there is $u \in V(P_e) \setminus \{r\}$ s. t. $L_1(u) = v$;
2. the number of leaves of T is $|\mathcal{E}|$.





Example

Simplicial complex $\mathcal{K}$ composed of eight vertices $\{v_1, \ldots, v_8\}$.

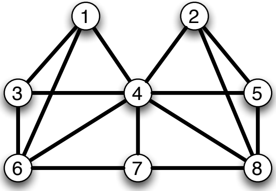


Maximal simplices :

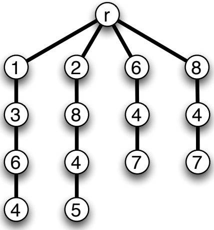
- tetrahedron induced by $\{v_1, v_3, v_4, v_6\}$,
- tetrahedron induced by $\{v_2, v_4, v_5, v_8\}$,
- triangle induced by $\{v_4, v_6, v_7\}$,
- triangle induced by $\{v_4, v_7, v_8\}$.

Example

Simplicial complex $\mathcal{K}$



Tree representation T composed of 15 vertices



Maximal simplices $\Leftrightarrow$ paths between r and the leafs :

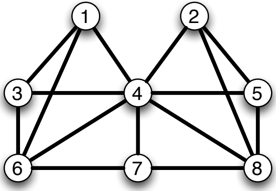
- tetrahedron induced by $\{v_1, v_3, v_4, v_6\}$ $\Leftrightarrow$ path $(r, 1, 3, 4, 6)$,
- tetrahedron induced by $\{v_2, v_4, v_5, v_8\}$ $\Leftrightarrow$ path $(r, 2, 4, 5, 8)$,
- triangle induced by $\{v_4, v_6, v_7\}$ $\Leftrightarrow$ path $(r, 6, 4, 7)$,
- triangle induced by $\{v_4, v_7, v_8\}$ $\Leftrightarrow$ path $(r, 8, 4, 7)$.

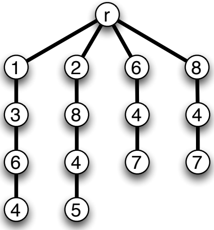
Tree representation

Definition

Let $\mathcal{K}$ be a simplicial complex and let $\mathcal{E}$ be the set of maximal simplices of $\mathcal{K}$. A tree $T = (V, E, L_1)$ rooted at $r \in V$, $L_1 : V \to \mathcal{V}$, is a tree representation of $\mathcal{K}$ if and only if

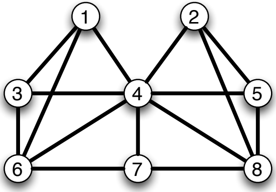
1. $\forall e \in \mathcal{E}$, there is a simple path P_e in T between r and a leaf s. t. $\forall v \in e$, there is $u \in V(P_e) \setminus \{r\}$ s. t. $L_1(u) = v$;
2. the number of leaves of T is $|\mathcal{E}|$.



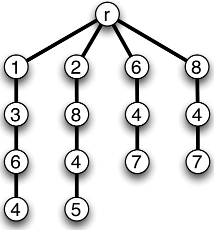


Example

Simplicial complex $\mathcal{K}$



Tree representation T composed of 15 vertices



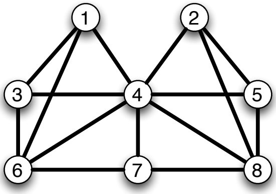
Maximal simplices $\Leftrightarrow$ paths between r and the leafs :

- tetrahedron induced by $\{v_1, v_3, v_4, v_6\}$ $\Leftrightarrow$ path $(r, 1, 3, 4, 6)$,
- tetrahedron induced by $\{v_2, v_4, v_5, v_8\}$ $\Leftrightarrow$ path $(r, 2, 4, 5, 8)$,
- triangle induced by $\{v_4, v_6, v_7\}$ $\Leftrightarrow$ path $(r, 6, 4, 7)$,
- triangle induced by $\{v_4, v_7, v_8\}$ $\Leftrightarrow$ path $(r, 8, 4, 7)$.

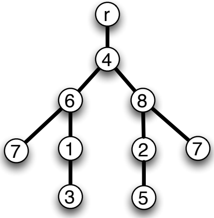
Tree representation T of $\mathcal{K}$ composed of < 15 vertices ?

Example

Simplicial complex $\mathcal{K}$



Tree representation T
composed of 10 vertices

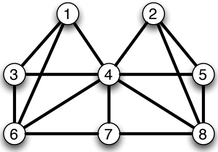


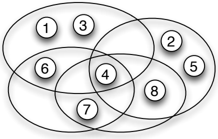
Maximal simplices $\Leftrightarrow$ paths between r and the leafs :

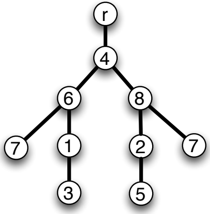
- tetrahedron induced by $\{v_1, v_3, v_4, v_6\} \Leftrightarrow$ path $(r, 4, 6, 1, 3)$,
- tetrahedron induced by $\{v_2, v_4, v_5, v_8\} \Leftrightarrow$ path $(r, 4, 8, 2, 5)$,
- triangle induced by $\{v_4, v_6, v_7\} \Leftrightarrow$ path $(r, 4, 6, 7)$,
- triangle induced by $\{v_4, v_7, v_8\} \Leftrightarrow$ path $(r, 4, 8, 7)$.

Simplicial complexes modeled by hypergraphs

Simplicial complex $\mathcal{K}$ | Hypergraph $\mathcal{H}$ | Tree representation T



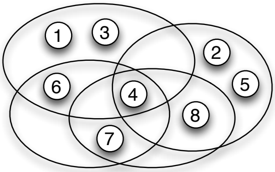


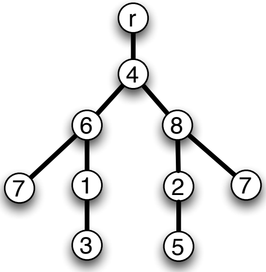


Definition

Let $\mathcal{H} = (\mathcal{V}, \mathcal{E})$ be a hypergraph. A tree $T = (V, E, L_1)$ rooted at $r \in V$, $L_1 : V \to \mathcal{V}$, is a tree representation of $\mathcal{H}$ if and only if

1. $\forall e \in \mathcal{E}$, there is a simple path P_e in T between r and a leaf s. t. $\forall v \in e$, there is $u \in V(P_e) \setminus \{r\}$ s. t. $L_1(u) = v$;
2. the number of leaves of T is $|\mathcal{E}|$.



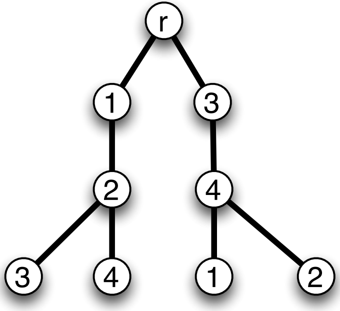


Need for additional constraints

Let $\mathcal{H} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} = \{v_1, v_2, v_3, v_4\}$ and
$\mathcal{E} = \{\{v_1, v_2, v_3\}, \{v_1, v_2, v_4\}, \{v_1, v_3, v_4\}, \{v_2, v_3, v_4\}$.

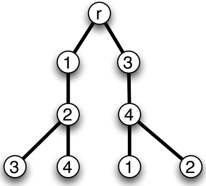
Deciding if $\{v_1, v_2, v_3\}$ is in T ?

$\rightarrow$ We do not know which neighbor of r to choose.



Deciding if $\{v_1, v_2, v_3\}$ is in T ?

$\rightarrow$ We do not know which neighbor of r to choose.



Lemma (simplex search algorithm for tree representation)

Let **A** *be any algorithm for the problem of searching a given maximal simplex. Then, there exists a hypergraph* $\mathcal{H} = (\mathcal{V}, \mathcal{E})$ *and a tree representation* T *of* $\mathcal{H}$ *such that the time complexity of* **A** *is* $\Omega(|\mathcal{V}|^2) = \Omega(|V(T)|)$.

Global tree representation

Definition

Let $\mathcal{H} = (\mathcal{V}, \mathcal{E})$ be a hypergraph. A tree $T = (V, E, L_1, L_2)$ rooted at $r \in V$, $L_1 : V \to \mathcal{V}$, $L_2 : V \to [\![0, |\mathcal{V}|]\!]$, is a global tree representation of $\mathcal{H}$ if and only if

1. $\forall e \in \mathcal{E}$, there is a simple path P_e in T between r and a leaf s. t. $\forall v \in e$, there is $u \in V(P_e) \setminus \{r\}$ s. t. $L_1(u) = v$;
2. the number of leaves of T is $|\mathcal{E}|$;
3. $\forall u, u' \in V$, then $L_1(u) = L_1(u') \Leftrightarrow L_2(u) = L_2(u')$;
4. for every simple path $P = (r, u_1, \ldots, u_t)$ of T, then $L_2(u_i) < L_2(u_{i+1})$ for all $i \in [\![1, t-1]\!]$.

$\Rightarrow$ Every path of T follows a global order σ induced by L_2.

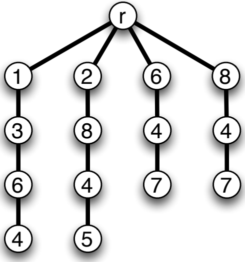
Global tree representation

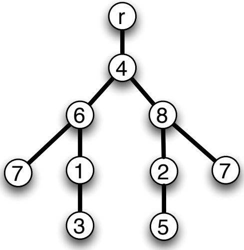
$$\mathcal{E} = \{\{v_1, v_3, v_4, v_6\}, \{v_2, v_4, v_5, v_8\}, \{v_4, v_6, v_7\}, \{v_4, v_7, v_8\}.$$

Every path of T follows a global order σ induced by L_2.

$\sigma = (v_1, v_2, v_3, v_6, v_8, v_4, v_5, v_7)$

$\sigma = (v_4, v_6, v_8, v_7, v_1, v_3, v_2, v_5)$

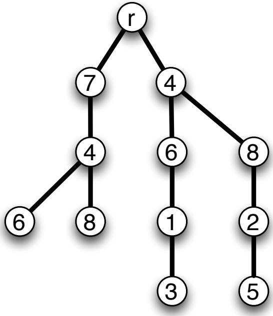




Tree obtained with $\sigma = (v_7, v_4, v_6, v_8, v_1, v_3, v_2, v_5)$?

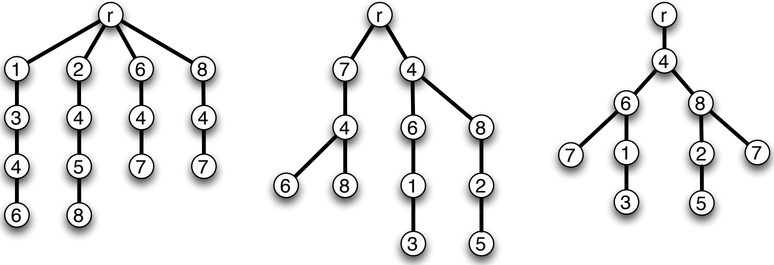
$$\mathcal{E} = \{\{v_1, v_3, v_4, v_6\}, \{v_2, v_4, v_5, v_8\}, \{v_4, v_6, v_7\}, \{v_4, v_7, v_8\}.$$

Global tree representation obtained with
$\sigma = (v_7, v_4, v_6, v_8, v_1, v_3, v_2, v_5)$:



Different global tree representations

$$\mathcal{E} = \{\{v_1, v_3, v_4, v_6\}, \{v_2, v_4, v_5, v_8\}, \{v_4, v_6, v_7\}, \{v_4, v_7, v_8\}.$$



Complexity of search algorithms

Let $\mathcal{H}$ be any hypergraph.
Let T be a global tree representation of $\mathcal{H}$.

Complexity of an algorithm for the problem of searching a given maximal simplex ?

(the set of neighbors of every node of T is ordered according to L_2)

$d_{\mathcal{H}}$: dimension of $\mathcal{H}$.
$|\mathcal{E}|$: number of hyperedges of $\mathcal{H}$.
$|V(T)|$: number of nodes of T.
Δ_T : maximum degree of T.
...

Complexity of search algorithms

- $O(\log_2(\Delta_T))$-time for dertermining the "neighbor" corresponding to the current node of the simplex.
- Each path between the root and a leaf has size at most $d_{\mathcal{H}}$.

Lemma (simplex search algorithm for global tree representation)

Let $\mathcal{H}$ be any hypergraph and let T be a global tree representation of $\mathcal{H}$. There exists a $O(d_{\mathcal{H}} \log_2(\Delta_T))$-time complexity algorithm for the problem of searching a given maximal simplex.

$d_{\mathcal{H}}$: dimension of $\mathcal{H}$.
Δ_T : maximum degree of T.

Complexity of search algorithms

Lemma (simplex search algorithm for tree representation)

Let **A** *be any algorithm for the problem of searching a given maximal simplex. Then, there exists a hypergraph* $\mathcal{H} = (\mathcal{V}, \mathcal{E})$ *and a tree representation* T *of* $\mathcal{H}$ *such that the time complexity of* **A** *is* $\Omega(|\mathcal{V}|^2) = \Omega(|V(T)|)$.

Lemma (simplex search algorithm for global tree representation)

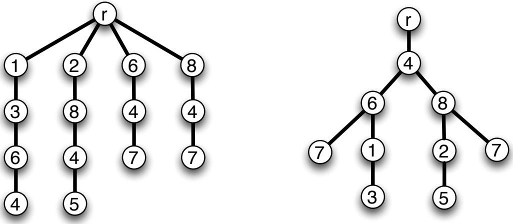
Let $\mathcal{H}$ *be any hypergraph and let* T *be a global tree representation of* $\mathcal{H}$*. There exists a* $O(d_{\mathcal{H}} \log_2(\Delta_T))$*-time complexity algorithm for the problem of searching a given maximal simplex.*

$d_{\mathcal{H}}$: dimension of $\mathcal{H}$.
Δ_T : maximum degree of T.

Problem

Let $\mathcal{H} = (\mathcal{V}, \mathcal{E})$ be any hypergraph. The global tree representation problem consists in computing the maximum max^*_{global} such that there exists a global tree representation T^*_{global} of $\mathcal{H}$ with

$$|V(T^*_{global})| = -max^*_{global} + 1 + \sum_{e \in \mathcal{E}} |e|.$$



$\max^*_{global} = 5$

Complexity results

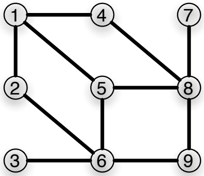
Theorem

The decision variant of the global tree representation problem is NP-complete even for the class of cubic graphs and for the class of planar graphs of degree at most three.

Idea of the proof for graphs ? Reduction with vertex cover problem ?

A vertex cover X of G is s. t. $\forall \{u, v\} \in \mathcal{E}$, then $\{u, v\} \cap X \neq \emptyset$.

Optimal global tree representation for :



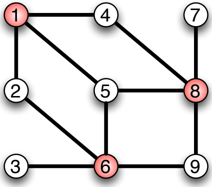
Sketch of the proof

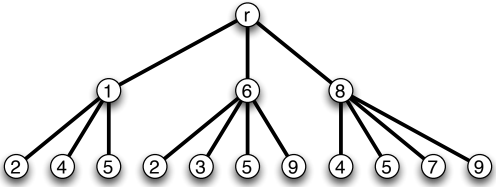
Let $\mathcal{H} = (\mathcal{V}, \mathcal{E})$ be any graph. The number of leafs of any tree representation is the number of edges $|\mathcal{E}|$.

$\Rightarrow$ The global tree representation problem is equivalent to minmize the number of neighbors of the root r.

The set of neighbors of r forms a vertex cover X of G ($\forall \{u, v\} \in \mathcal{E}$, then $\{u, v\} \cap X \neq \emptyset$).

Vertex cover problem is NP-complete, and so global tree representation problem is NP-complete.





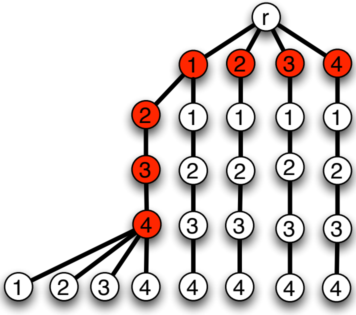
Let $\mathcal{H} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} = \{a_1, a_2, a_3, a_4, b_1, b_2, b_3, b_4\}$ and $\mathcal{E} = \{\{a_1, a_2, a_3, a_4, b_1\}, \{a_1, a_2, a_3, a_4, b_2\}, \{a_1, a_2, a_3, a_4, b_3\}, \{a_1, a_2, a_3, a_4, b_4\}, \{b_1, b_2, b_3, b_4, a_1\}, \{b_1, b_2, b_3, b_4, a_2\}, \{b_1, b_2, b_3, b_4, a_3\}, \{b_1, b_2, b_3, b_4, a_4\}\}$.

Find an optimal global tree representation of $\mathcal{H}$?

In other words, find an optimal ordering for the global tree representation problem ?

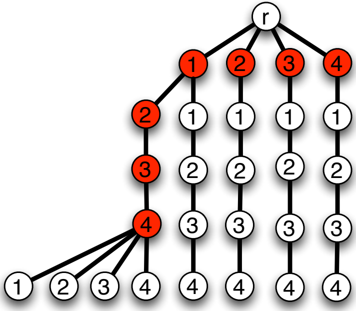
$$\mathcal{E} = \{\{a_1, a_2, a_3, a_4, b_1\}, \{a_1, a_2, a_3, a_4, b_2\}, \{a_1, a_2, a_3, a_4, b_3\}, \{a_1, a_2, a_3, a_4, b_4\}, \{b_1, b_2, b_3, b_4, a_1\}, \{b_1, b_2, b_3, b_4, a_2\}, \{b_1, b_2, b_3, b_4, a_3\}, \{b_1, b_2, b_3, b_4, a_4\}\}.$$

$$\sigma = (a_1, a_2, a_3, a_4, b_1, b_2, b_3, b_4)$$



Limit of global tree representation

$$\sigma = (a_1, a_2, a_3, a_4, b_1, b_2, b_3, b_4)$$



$\Rightarrow$ We cannot factorize $\{b_1, b_2, b_3, b_4\}$.

$\Rightarrow$ Local tree representation.

Local tree representation (recursive definition)

Definition

Let $\mathcal{H} = (\mathcal{V}, \mathcal{E})$ be a hypergraph. A local tree representation of $\mathcal{H}$ is a node-labeled tree $T = (V, E, L_1, L_2)$ rooted at $r \in V$, $L_1 : V \rightarrow \mathcal{V}$, $L_2 : V \rightarrow [\![0, |\mathcal{V}|]\!]$, such that :

1. if $|\mathcal{E}| = 0$, then $T = (\{r\}, \emptyset)$;
2. if $|\mathcal{E}| \geq 1$, then there exists a node $u \in N_T(r)$, with $L_1(u) = v \in \mathcal{V}$, such that :
   1. for every $u' \in N_T(r) \setminus \{u\}$, then $L_2(u) < L_2(u')$;
   2. the tree $T[u]$ rooted at $r' = u$ is a local tree representation of $(\mathcal{V} \setminus \{v\}, \mathcal{E}_v)$;
   3. the tree $T \setminus T[u]$ rooted at r is a local tree representation of $(\mathcal{V} \setminus \{v\}, \bar{\mathcal{E}}_v)$.

$\mathcal{E}_v = \{e \setminus \{v\} \mid v \in e, e \in \mathcal{E}\}$.
$\bar{\mathcal{E}}_v = \{e \mid v \notin e, e \in \mathcal{E}\}$.

Local tree representation problem

Problem

Let $\mathcal{H} = (\mathcal{V}, \mathcal{E})$ be any hypergraph. The local tree representation problem consists in computing the maximum max^*_{local} such that there exists a local tree representation T^*_{local} of $\mathcal{H}$ with $|V(T^*_{local})| = -max^*_{local} + 1 + \sum_{e \in \mathcal{E}} |e|$.

Example

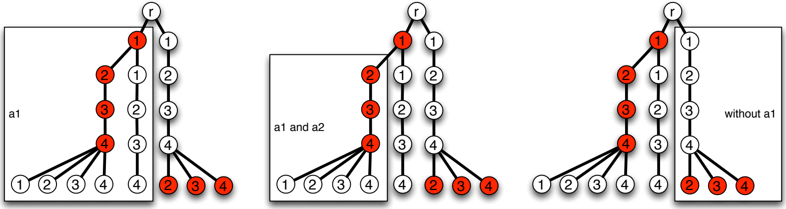
Definition

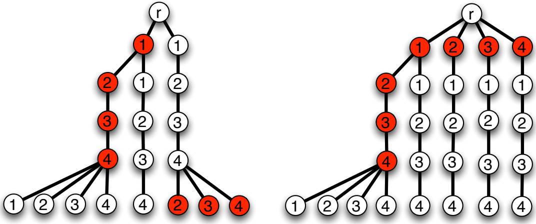
Let $\mathcal{H} = (\mathcal{V}, \mathcal{E})$ be a hypergraph. A local tree representation of $\mathcal{H}$ is a node-labeled tree $T = (V, E, L_1, L_2)$ rooted at $r \in V$, $L_1 : V \to \mathcal{V}$, $L_2 : V \to [\![0, |\mathcal{V}|]\!]$, such that :

1. if $|\mathcal{E}| = 0$, then $T = (\{r\}, \emptyset)$;
2. if $|\mathcal{E}| \geq 1$, then there exists a node $u \in N_T(r)$, with $L_1(u) = v \in \mathcal{V}$, such that :
   1. for every $u' \in N_T(r) \setminus \{u\}$, then $L_2(u) < L_2(u')$;
   2. the tree $T[u]$ rooted at $r' = u$ is a local tree representation of $(\mathcal{V} \setminus \{v\}, \mathcal{E}_v)$;
   3. the tree $T \setminus T[u]$ rooted at r is a local tree representation of $(\mathcal{V} \setminus \{v\}, \bar{\mathcal{E}}_v)$.

Optimal local tree representation for $\mathcal{E} = \{\{a_1, a_2, a_3, a_4, b_1\}, \{a_1, a_2, a_3, a_4, b_2\}, \{a_1, a_2, a_3, a_4, b_3\}, \{a_1, a_2, a_3, a_4, b_4\}, \{b_1, b_2, b_3, b_4, a_1\}, \{b_1, b_2, b_3, b_4, a_2\}, \{b_1, b_2, b_3, b_4, a_3\}, \{b_1, b_2, b_3, b_4, a_4\}\}$?

Example





Property

Let $\mathcal{H}$ be any hypergraph. Then, $|V(T^*_{global})| \geq |V(T^*_{local})|$ and $max^*_{global} \leq max^*_{local}$.

Lemma

For any $n \geq 1$, there exists a hypergraph $\mathcal{H} = (\mathcal{V}, \mathcal{E})$, with $n = |\mathcal{V}|$, such that $|V(T^*_{local})| \leq 5n$ and $|V(T^*_{global})| \geq n^2$, and $max^*_{local} \geq 2n^2 - 3n + 1$ and $max^*_{global} \leq (n+1)^2$.

Local versus Global $\rightarrow$ Global for the search complexity

Lemma (simplex search algorithm for global tree representation)

Let $\mathcal{H}$ be any hypergraph and let T be a global tree representation of $\mathcal{H}$. There exists a $O(d_{\mathcal{H}} \log_2(\Delta_T))$-time complexity algorithm for the problem of searching a given maximal simplex.

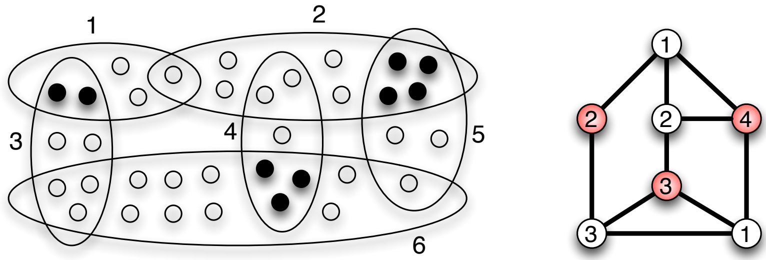
Lemma (simplex search algorithm for local tree representation)

Let $\mathcal{H}$ be any hypergraph and let T be a local tree representation of $\mathcal{H}$. There exists a $O(d^2_{\mathcal{H}} \log_2(\Delta_T))$-time complexity algorithm for the problem of searching a given maximal simplex.

Theorem

The global and local tree representation problems are in P for hypergraphs with maximum degree at most two.

Maximum weighted independent set of a line graph → polynomial



Tree representations of bounded degree hypergraphs

Theorem

The decision variants of the global and local tree representation problems are NP-complete even for hypergraphs with maximum degree three.

Corollary

The global and local tree representation problems cannot be approximated within $\frac{k}{\log(k)}$-ratio in polynomial time for the class of hypergraphs of maximum degree k, where $k \geq 3$ is a constant integer, unless $P = NP$.

Corollary

The global and local tree representation problems admit a polynomial time $\frac{k}{2}$-approximation algorithm for the class of hypergraphs of maximum degree k, where $k \geq 3$ is a constant integer.

Conclusion

Tree representation

- Decreasing the size of the representation.
- Updating the structure efficiently.

Global representation versus Local representation

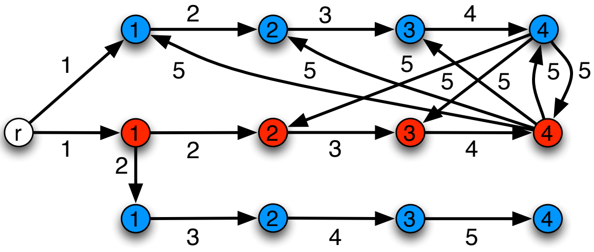
- Global representation : efficient in terms of time complexity for searching a given simplex.
- Local representation : efficient in terms of size of the structure

Complexity results

- NP-complete for graphs.
- Hard to approximate for bounded degree hypergraphs.

Future works

- Representation of all the simplexes (not only maximal).
- Digraph representation : efficient in terms of size of the structure and time complexity for searching a simplex.



Internship at Inria Sophia Antipolis Méditerranée

Directed graph representation of simplicial complexes.

- Definition of the structure
- Complexity results : hardness results...
- Algorithms : Branch and Bound, approximation...
- Implementation

Contact
Jean-Daniel.Boissonnat@inria.fr, Dorian.Mazauric@inria.fr
(project-team Geometrica)

References

J.-D. Boissonnat, C. S. Karthik, and S. Tavenas.
Building efficient and compact, data structures for simplicial complexes.
Submitted to SoCG 2015.

J.-D. Boissonnat and C. Maria.
The simplex tree : An efficient data structure for general simplicial complexes.
Algorithmica, 70(3) :406–427, 2014.

J.-D. Boissonnat and D. Mazauric.
On the complexity of the representation of simplicial complexes by trees.
Submitted to SoCG 2015.