

Ecole thématique « Mathématiques et philosophie contemporaines, XII »

Saint-Ferréol, 26 juin 2025

Dites-le avec des sphères !

Etienne Fieux

IMT, Université de Toulouse

plan

1. Etude de $P \stackrel{?}{=} NP$ via 2-SAT et 3-SAT
2. Approche « géométrique »
3. Quelques calculs

résumé

En 2008, dans l'article *Boolean formulae, hypergraphs and combinatorial topology* (<https://arxiv.org/abs/0808.0739>), les mathématiciens J. Conant et O. Thistlethwaite proposent une approche originale pour traiter le problème $P \stackrel{?}{=} NP$ en transformant une question de complexité algorithmique en une question de « géométrie » et de calculs de types d'homotopie.

Jusqu'ici, cette approche semble être restée lettre morte, mais le but de cet exposé est de présenter les idées qui la mènent, comme illustration des méthodes de la combinatoire topologique.

En parlant de « combinatoire topologique », on veut insister sur le fait que le propos est de résoudre un problème combinatoire (ou un problème de nature « discrète ») en s'aidant des outils classiques de la topologie algébrique (anciennement appelée « topologie combinatoire », pour dire l'idée d'utiliser la combinatoire pour résoudre des problèmes de topologie, au sens d'étude d'objets de nature « continue »). La résolution par Lovasz de la conjecture de Kneser en 1978 est souvent considérée comme résultat fondateur de la combinatoire topologique (https://en.wikipedia.org/wiki/Topological_combinatorics).

Complexité algorithmique :

P/NP comme 2-SAT/3-SAT

problème SAT : savoir si une formule de la logique propositionnelle est satisfiable ; *i.e.*, peut-on assigner des valeurs V ou F aux variables de sorte que la formule soit vraie ?

Théorème de Cook (1971)

SAT est NP-complet

Complexité exponentielle

Pour évaluer une proposition utilisant n variables, il « suffit » de dresser sa table de vérité.

Pour bien représenter ce que signifie un coût exponentiel, si $n = 80$, il faut remplir 2^{80} lignes pour calculer la table de vérité complète.

« S'il nous faut une microseconde pour traiter chaque ligne, alors il faudra 2^{80} microsecondes pour remplir toute la table, et 2^{80} microsecondes est de l'ordre de 38 milliards d'années. En comparaison, l'âge de l'Univers est d'environ 15 milliards d'années. ».

A MATHEMATICAL INTRODUCTION TO LOGIC, H. Enderton

extrait de LOGIQUE : FONDEMENTS ET APPLICATIONS., P. Le Barbenchon, S. Pinchinat, F. Schwarzenruber.

Les formules FNC et le problème k -SAT

Logique propositionnelle : NON (\neg), ET (\wedge), OU (\vee)

Littéraux : x, \bar{y}, z, \dots (avec $\bar{y} = \neg y$)

Clause : disjonction de littéraux : $x \vee y \vee \bar{z} \vee t$

Formule sous forme normale conjonctive (FNC) : conjonction de clauses

- $x \wedge \bar{x}$ est non satisfiable
- $(x \vee \bar{y} \vee \bar{z}) \wedge (x \vee \bar{y}) \wedge (\bar{x} \vee \bar{z})$ est satisfiable (avec $x : V$ et $y : F$)

Problème k -SAT : problème SAT pour les formules k -FNC, i. e. les formules FNC formées de clauses comprenant k littéraux.

- $(x \vee z) \wedge (y \vee \bar{x})$ est une formule 2-FNC.
- $(x \vee \bar{y} \vee \bar{z}) \wedge (x \vee \bar{y} \vee z) \wedge (y \vee \bar{x} \vee \bar{z})$ est une formule 3-FNC.

Théorème

3-SAT est NP-complet mais 2-SAT est P.

L'idée directrice

Thus an attempt to understand the P/NP question is to understand the difference between 2-SAT and 3-SAT.

J. Conant, O. Thistlethwaite

L'idée directrice est donc :

toute distinction entre 2-SAT et 3-SAT *pourrait* constituer une distinction entre P et NP .

Et la méthode sera :

pour tout nombre n de variables fixé, associer à 2-SAT (resp. 3-SAT) un complexe simplicial $|2 - sat - n|$ (resp. $|3 - sat - n|$) en espérant faire apparaître des distinctions « utiles » entre les groupes d'homotopie de ces deux espaces.

On peut se ramener à 3-SAT

A toute formule f CNF, on peut associer une formule CNF $z(f)$ dans 3-SAT « équisatisfiable » (i. e. f satisfiable $\iff z(f)$ satisfiable).

Par exemple :

f	$z(f)$
a	$(a \vee x \vee y) \wedge (a \vee \bar{x} \vee y) \wedge (a \vee x \vee \bar{y}) \wedge (a \vee \bar{x} \vee \bar{y})$
$a \vee b$	$(a \vee b \vee x) \wedge (a \vee b \vee \bar{x})$
$a \vee b \vee c \vee d$	$(a \vee b \vee x) \wedge (\bar{x} \vee c \vee d)$

nota bene : f et $z(f)$ ne sont pas des formules équivalentes !

SAT-n vu au travers de l'hypercube $H_n := [0, 1]^n$

Les sommets de H_n s'identifie aux valuations possibles (pour n variables x_1, x_2, \dots, x_n) :

$$\{\tau : \{1, 2, \dots, n\} \rightarrow \{0, 1\}\}$$

Pour toute clause $C = x_{i_1} \vee x_{i_2} \cdots \vee x_{i_k}$, toutes les valuations qui envoient $x_{i_1}, x_{i_2}, \dots, x_{i_k}$ sur 0 ne satisfont pas C .

Cela définit une $n - k$ -face de H_n (de *non satisfaction*!)

k-Codim Cube Noncover (Welzl)

Input : $n \in \mathbb{N}$ et un ensemble de faces $\{f_i\}$ de $\dim \geq n - k$.

Output : - Oui s'il existe un sommet du cube qui n'est pas dans $\bigcup_i f_i$
 - Non, sinon

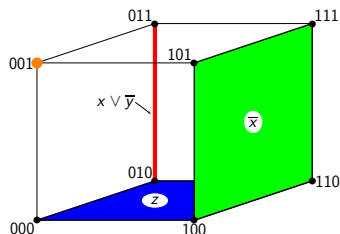
Deux illustrations

La formule CNF

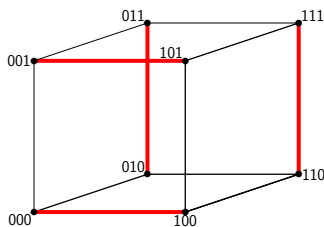
$$f = \bar{x} \wedge z \wedge (x \vee \bar{y})$$

est satisfiable

(vraie pour $\tau = 001$)



Pour toute formule formée avec 3 variables, il existe une formule équivalente formée d'au plus 4 clauses.



Version \ll géométrie \gg

En commençant par un aparté, autour du rôle joué par les ensembles partiellement ordonnées (« posets ») et les complexes simpliciaux dans ce passage entre objets « discrets » et objets « continus ».

Complexes simpliciaux (abstraits)

$K = (V(K), \Sigma(K))$ complexe simplicial (abstrait) :

- $V(K)$: ensemble (les sommets de K)
- $\Sigma(K)$: ensemble de sous-ensembles $\sigma = \{x_1, x_2, \dots, x_n\}$ de $V(K)$ (appelés *simplexes*) tel que $\tau \subset \sigma$ et $\sigma \in \Sigma(K) \Rightarrow \tau \in \Sigma(K)$.

Simplexes géométriques ou abstraits ?

0-simplexe

$\{a\}$



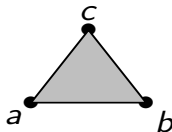
1-simplexe

$\{a, b\}$



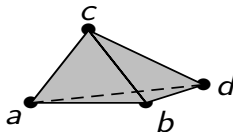
2-simplexe

$\{a, b, c\}$



3-simplexe

$\{a, b, c, d\}$

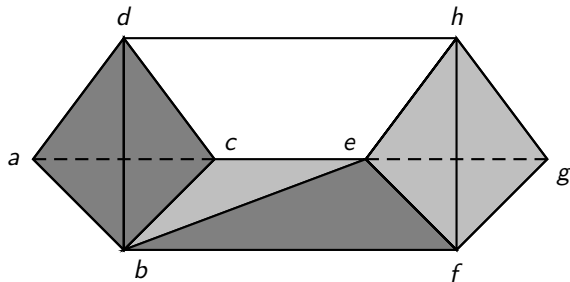


Un exemple de complexe simplicial K

$$V(K) = \{a, b, c, d, e, f, g, h\}$$

$\Sigma(K)$ est l'ensemble des simplexes :

- 0-simplexes : a, b, c, d, e, f, g, h
- 1-simplexes : $ab, ac, ad, bc, bd, cd, ce, be, bf, ef, eg, eh, fg, fh, gh$
- 2-simplexes : $abc, abd, acd, bcd, efg, efh, egh, fgh$
- 3-simplexe : $abcd$



$$c_0 = 8$$

$$c_1 = 16$$

$$c_2 = 10$$

$$c_3 = 1$$

Espaces topologiques finis

Un **ensemble ordonné** P (i.e. un ensemble P muni d'une relation d'ordre \leq : réflexive, antisymétrique et transitive) est appelé **poset** (partially ordered set).

Proposition

Soit X un ensemble fini ;

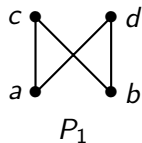
« munir X d'une topologie T_0 »

équivaut à

« munir X d'une relation d'ordre partiel »

- si X est T_0 , alors $x \leq y$ si $U_x \subset U_y$
- si (X, \leq) est un poset, alors $U_x := \{z \in X, z \leq x\}$

Posets : 2 exemples



ordre partiel

$$a, b \leq c, d$$

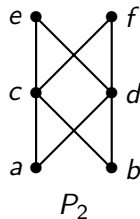
ouverts

$$U_a = \{a\}$$

$$U_b = \{b\}$$

$$U_c = \{a, b, c\}$$

$$U_d = \{a, b, d\}$$



ordre partiel

$$a, b \leq c, d \leq e, f$$

ouverts

$$U_a = \{a\}$$

$$U_b = \{b\}$$

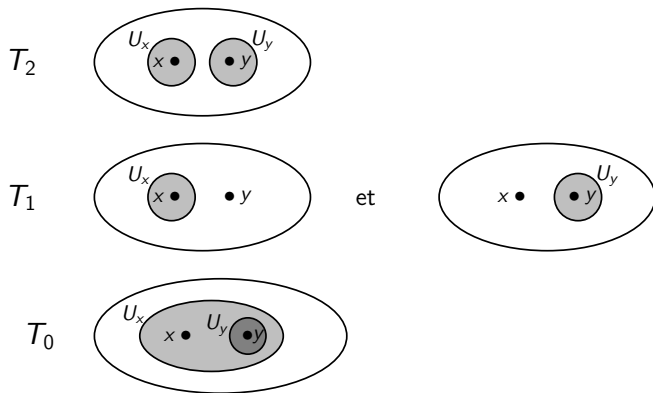
$$U_c = \{a, b, c\}$$

$$U_d = \{a, b, d\}$$

$$U_e = \{a, b, c, d, e\}$$

$$U_f = \{a, b, c, d, f\}$$

Topologies T_0 , T_1 et T_2 : $T_0 \subset T_1 \subset T_2$



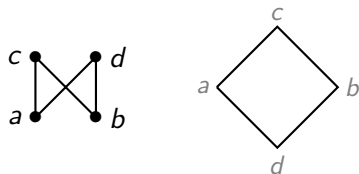
Un espace topologique muni d'une topologie T_2 est dit « de Hausdorff » (ou encore, espace séparé).

Un ensemble fini muni de la topologie T_2 est forcément « discret ». Mais la considération des topologies T_0 (c'est-à-dire, des posets) va permettre de voir que les espaces finis peuvent être très riches du point de vue de l'homotopie.

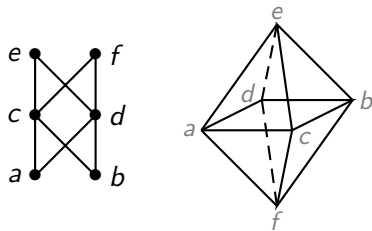
Des posets vers les complexes simpliciaux, 1

Si P est un poset, $\Delta(P)$ est le complexe simplicial défini par :

- P est l'ensemble de sommets
- les simplexes sont les $\{x_1, x_2, \dots, x_n\}$ pour chaque $x_1 < x_2 < \dots < x_n$



P_1 et $\Delta_{\mathcal{P}}(P_1)$



P_2 et $\Delta_{\mathcal{P}}(P_2)$

On dit que $\Delta(P)$ est la **réalisation géométrique** du poset P .

Des posets vers les complexes simpliciaux, 2

On a donc une application (un foncteur covariant...)

$$\Delta : \mathcal{P} \longrightarrow \mathcal{K}$$

et, pour un poset P donné, on définit

$$\psi_P : |\Delta(P)| \longrightarrow P$$

par $\psi(t) = x_0$ si t est dans le simplexe donné par $x_0 < x_1 < \dots < x_n$.

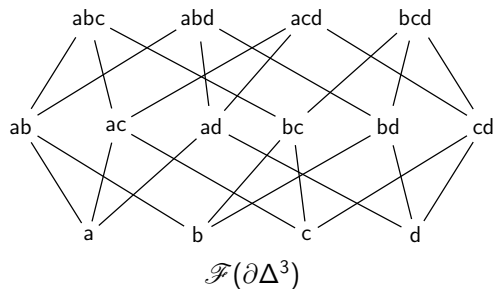
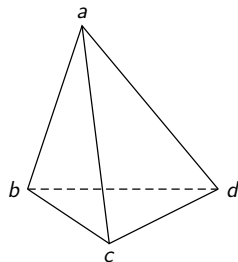
Théorème (McCord, 1966)

ψ_P est une équivalence faible d'homotopie

($\implies P$ et $\Delta(P)$ ont mêmes groupes d'homotopie).

Des complexes simpliciaux vers les posets, 1

Si K est un complexe simplicial, on lui associe le poset $\mathcal{F}(K)$, donné par l'ensemble des simplexes de K muni de la relation d'inclusion.



$\partial\Delta^3$ est le bord du tétraèdre plein, soit encore le « tétraèdre vide », composé des sommets $\{a\}, \{b\}, \{c\}, \{d\}$, des arêtes $\{a, b\}, \{a, c\}, \{a, d\}, \{b, c\}, \{b, d\}, \{c, d\}$ et des triangles pleins $\{a, b, c\}, \{a, b, d\}, \{a, c, d\}$ et $\{b, c, d\}$.

On dit que $\mathcal{F}(K)$ est le « poset des faces » du complexe simplicial K .

Des complexes simpliciaux vers les posets, 2

On a également une application (qui est aussi un foncteur covariant...)

$$\mathcal{F} : \mathcal{K} \longrightarrow \mathcal{P}$$

Théorème (McCord, 1966)

$|K|$ et $\mathcal{F}(K)$ ont même type d'homotopie faible.

Plus précisément, on a une équivalence faible d'homotopie

$$\phi_K : |K| \sim |Sd(K)| = |\Delta \circ \mathcal{F}(K)| \longrightarrow \mathcal{F}(K)$$

qui utilise $\Delta \circ \mathcal{F} = Sd$ (où Sd désigne la subdivision barycentrique) et le théorème précédent : $\phi_K = \psi_{\mathcal{F}(K)}$.

Retour à 2-SAT vs 3-SAT

Le complexe simplicial $|k - \text{SAT} - n|$

J. Conant et O. Thistlethwaite, 2010

- $V(|k - \text{SAT} - n|) = k - \text{SAT} - n$, l'ensemble des formules CNF formées uniquement de k -clauses et satisfiables
- $\{f_0, f_1, \dots, f_k\}$ est un k -simplexe si $f_0 \Rightarrow f_1 \Rightarrow \dots \Rightarrow f_k$ (en sous-entendant que chaque implication est vraie)

Théorème

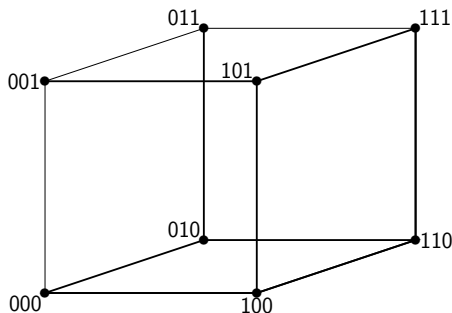
$|k - \text{SAT} - n|$ a la même type d'homotopie que $\Theta(\text{Cube}(n, n - k))$.

où $\Theta(\text{Cube}(n, n - k))$ est le complexe simplicial dont :

- les sommets sont ceux de l'hypercube H_n , soit $\{0, 1\}^n$
- et les simplexes sont les sous-ensembles des complémentaires des $(n - k)$ -faces.

Illustration avec $Cube(3, 1)$

$Cube(3, 1)$ est associé à $|2 - sat - 3|$:



Le complexe simplicial $\Theta(Cube(3, 1))$ a 12 facettes (ou simplexes maximaux pour l'inclusion) : chaque arête ab définissant la facette $\{0, 1\}^3 \setminus \{a, b\}$.

Un mot sur la preuve

On considère le recouvrement de $|k - \text{SAT} - n|$:

$$\mathcal{U} = \{U_\tau, \tau \in \text{Val}_n = \{0, 1\}^{\{1, 2, \dots, n\}}\}$$

Pour $\tau = (i_1, i_2, \dots, i_n)$, U_τ est l'ensemble des simplexes $\{f_0, f_1, \dots, f_k\}$ pour lesquels chaque formule f_i est satisfaite par τ ($i_k = 0$ signifie « x_k est F » et $i_k = 1$ signifie « x_k est V »).

Et on vérifie ensuite :

- $\mathcal{N}(\mathcal{U}) \simeq |k - \text{SAT} - n|$ (même type d'homotopie)
- $\mathcal{N}(\mathcal{U}) \cong \Theta(\text{Cube}(n, n - k))$ (complexes simpliciaux isomorphes)

La « cheville ouvrière » de cette preuve repose sur la notion de « nerf ».

L'importance des nerfs n'est pas nouvelle

Hausdorff's work opened up a whole new world to the point-set topologists. For combinatorialists, however, it was quite sufficient to consider triangulable spaces and at first it seemed that combinatorial methods could not be extended to more general classes of spaces. However an idea of Alexandroff's, dating from 1928, enabled combinatorial methods to be applied in great generality. As he put it in [] : « The transfer of homological objects to more general geometrical objects than polyhedra... based on my... concept of the nerve of a covering of a given space makes it possible to transfer to practically all topological spaces the fundamental concepts of combinatorial topology ». Remarkably it was this idea which led to Hausdorff developing a late interest in combinatorial topology.*

I. M. James, Combinatorial topology versus point-set topology, HANDBOOK OF THE HISTORY OF GENERAL TOPOLOGY, vol. 3, 809-834.

[*] Poincaré and topology, ICM Amsterdam, 1954 ; PSPM 39, part 2, 1983. Alexandroff ajoute en note : « Of course, the first sources of my concept of nerve lie in what Poincaré called "reciprocal polyhedron" »

La notion de nerf

Le **nerf** $\mathcal{N}(\mathcal{A})$ d'une famille d'ensembles $\mathcal{A} = \{A_1, A_2, \dots, A_n\}$ est le complexe simplicial formé

- de n sommets x_1, x_2, \dots, x_n
- et des simplexes $\sigma = \{x_{i_1}, x_{i_2}, \dots, x_{i_k}\}$ si $\bigcap_{j=1}^k A_{i_j} \neq \emptyset$

Soit $\mathcal{A} = \{A_1, A_2, \dots, A_n\}$ un recouvrement d'un espace topologique X par des ouverts (i.e. : chaque A_i est ouvert et $X = \bigcup_{i=1}^n A_i$).

Nerve theorem (Leray, 1945 ; Borsuk 1948 ; Weil, 1952 ; Mc Cord, 1966 ; Björner et al. 1985)

Si la famille (A_i) est localement finie et si toute intersection non vide $\bigcap_{j=1}^k A_{i_j}$ est contractile,
alors X et $\mathcal{N}(\mathcal{A})$ ont même type d'homotopie.

Pourquoi des sphères ?

L'idée de chercher à différencier les problèmes 2-SAT et 3-SAT va alors se traduire ici par la nécessité, pour tout entier $n \geq 2$, de différencier les espaces $|2 - \text{sat} - n|$ et $|3 - \text{sat} - n|$. On cherchera à faire cela au travers du calcul des groupes d'homotopie de ces espaces. Et c'est ce calcul qui fera apparaître de multiples bouquets de sphères.

Des bouquets de cercles

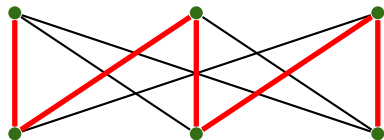
Type d'homotopie d'un graphe (où un graphe est vu comme un complexe simplicial de dimension 1). Cet exemple montre dans le cas le plus simple (dimension 1) comment des sphères (en dimension 1 : des cercles) apparaissent lorsqu'on calcule les groupes d'homotopie.

Théorème

Soit X un graphe connexe et T un arbre maximal dans X .

Alors le quotient X/T est un bouquet de cercles, avec autant de cercles que d'arêtes de X qui ne sont pas dans T .

De plus, la projection $X \rightarrow X/T$ est une équivalence d'homotopie.

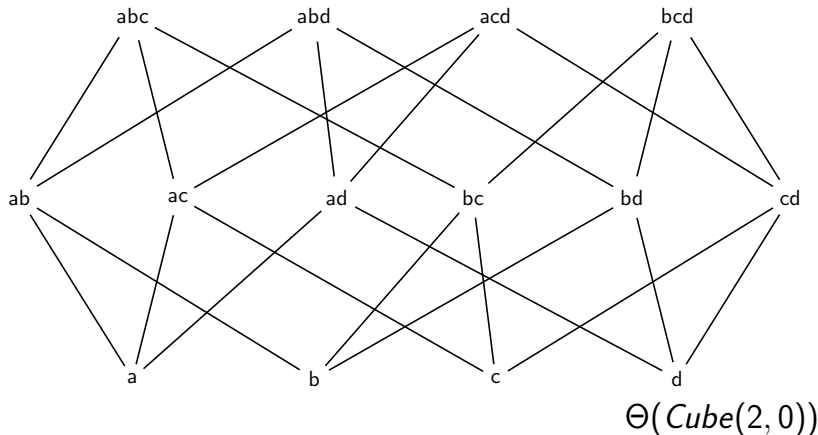


Exemple :

$$K_{3,3} \simeq \bigvee_{E(X) \setminus E(T)} S^1 = S^1 \vee S^1 \vee S^1 \vee S^1$$

Avec $n = 2$ variables : $|2 - sat - 2| \simeq S^2$

Avec $a = (0, 0) = \bar{x} \vee \bar{y}$, $b = (1, 0) = x \vee \bar{y}$, $c = (0, 1) = \bar{x} \vee y$ et $d = (1, 1) = x \vee y$, on voit que $\Theta(Cube(2, 0)) = \partial(\Delta^3) \simeq S^2$. :



Avec $n = 3$ variables :

$$|3 - \text{sat} - 3| \simeq S^6 \text{ et } |2 - \text{sat} - 3| \simeq S^4 \vee S^4 \vee S^4$$

- Comme pour $|2 - \text{sat} - 2| \simeq S^2$, on obtient $|3 - \text{sat} - 3| \simeq S^6$ en notant que $\Theta(\text{Cube}(3, 0)) = \partial(\Delta^7)$.
- Le résultat $|2 - \text{sat} - 3| \simeq S^4 \vee S^4 \vee S^4$ est moins direct mais peut s'obtenir en appliquant la théorie de Morse discrète.

Théorie de Morse discrète

Soit K un complexe simplicial abstrait.

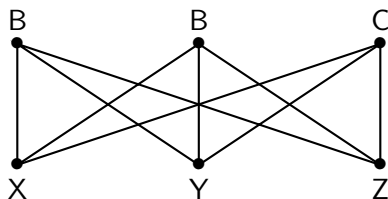
Théorème

S'il existe un matching acyclique sur un ensemble M de simplexes de K , alors il existe un CW-complexe \mathcal{X} tel que

- 1) en toute dimension d , le nombre de d -cellules de \mathcal{X} est égal au nombre de d -simplexes qui ne sont pas dans M
- 2) et $K \simeq \mathcal{X}$ (équivalence d'homotopie)

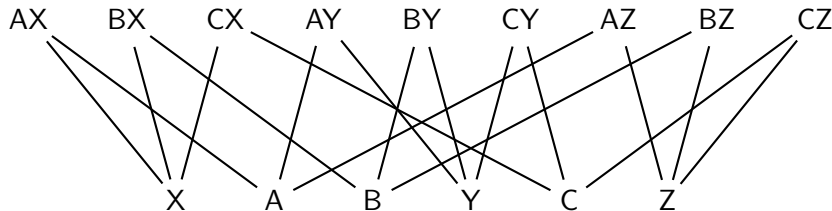
Théorie de Morse discrète, exemple 1 : $K_{3,3}$

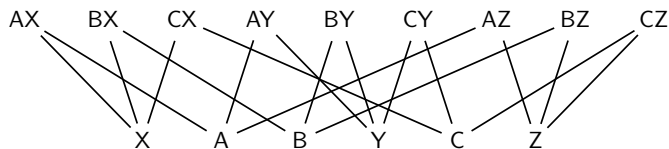
Le graphe $K_{3,3}$:



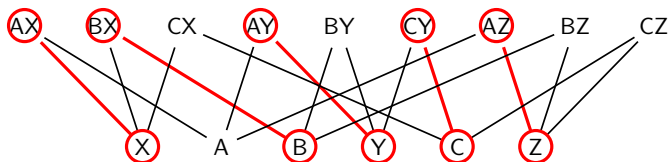
$$K_{3,3} \simeq \bigvee_{i=1}^4 S^1$$

On applique la théorie de Morse discrète au poset des faces de $K_{3,3}$ (vu comme complexe simplicial) :





On applique le matching : (X,AX) , (B,BX) , (Y,AY) , (C,CY) , (Z,AZ) :



Il reste 4 cellules de dimension 1 et une cellule de dimension 0 :

CX

BY

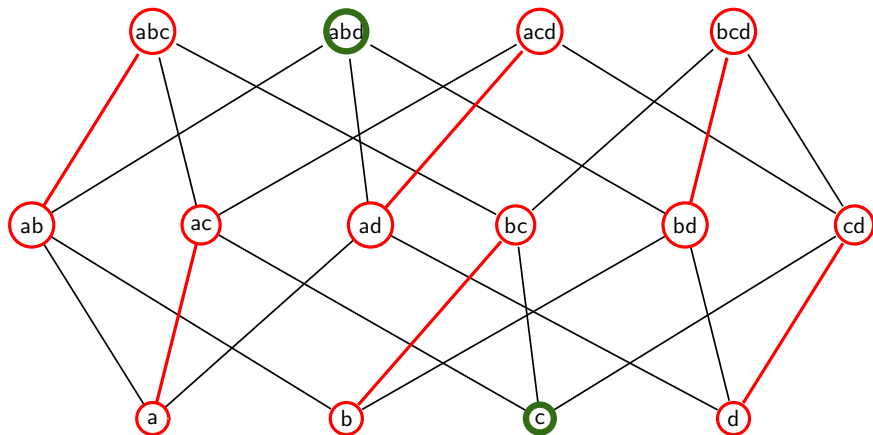
BZ

CZ

A

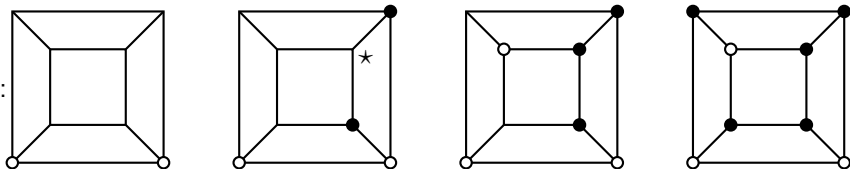
$$\text{d'où } K_{3,3} \simeq \bigvee_{i=1}^4 S^1$$

Théorie de Morse discrète, exemple 2 : $\Theta(Cube(2,0))$

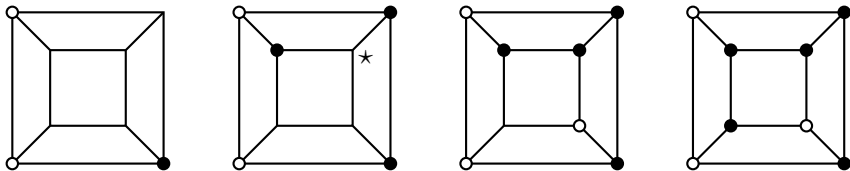


Exemple 3 : $\Theta(Cube(3,1)) \simeq S^4 \vee S^4 \vee S^4$

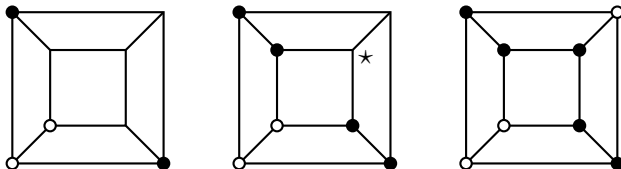
1er cas :



2e cas :



3e cas :

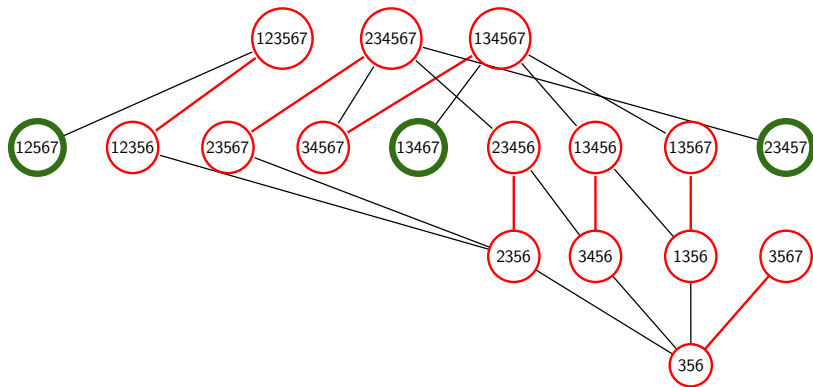


preuve de
Conant
et
Thistlethwaite

Autre visualisation de : $\Theta(Cube(3, 1)) \simeq \bigvee_{i=1}^3 S^4$

Par applications successives de la théorie de Morse discrète :

1. On conserve le sommet $\{0\}$ et on supprime tous les simplexes qui contiennent 0.
2. On conclut alors :



Retour à la question du départ : 2 conjectures

Le calcul du type d'homotopie de $\Theta(\text{Cube}(n, n - k))$ dans le cas général est difficile... et reste à faire !

Les deux conjectures qui suivent traduiraient la distinction recherchée entre 2-SAT et 3-SAT (ou k -SAT pour $k \geq 3$) :

Conjectures (J. Conant et O. Thistlethwaite, 2010)

1. $\Theta(\text{Cube}(n, n - 2))$ a le même type d'homotopie que $\bigvee_{(2n-3)!!} S^{2n-2}$.
2. Pour k -SAT- n , $k \geq 3$ et n assez grand, on obtient un bouquet de sphères qui ne sont pas toutes de même dimension.

P versus NP, la lutte reste acharnée...

<https://wscor.win.tue.nl/woeginger/P-versus-NP.htm>

