

# Analyse expérimentale des clauses apprises dans les solveurs modernes pour la satisfiabilité (SAT).

Adueni Adjoba Eunice AKANI

UE Communication

24 janvier 2019

Encadrant :  
Djamal HABET

- 1 Introduction
- 2 Définitions
- 3 Solveur CDCL
  - Principe
  - Apprentissage des clauses
- 4 Analyse et évaluation des clauses apprises
  - Environnement de test
  - Extraction des critères d'évaluation
  - Corrélation entre critères
  - Classification des clauses
- 5 Conclusion et perspective

# Introduction

- Le problème SAT est un problème de décision qui est au coeur de la théorie de la complexité.
- Plusieurs questions sont posées sur la résolution efficace du problème SAT
- Ce travail s'y intègre.

# Définitions

- **Variable booléenne** : Une variable à deux états (soit vrai ou faux)
- **Clause** : Une disjonction de littéraux

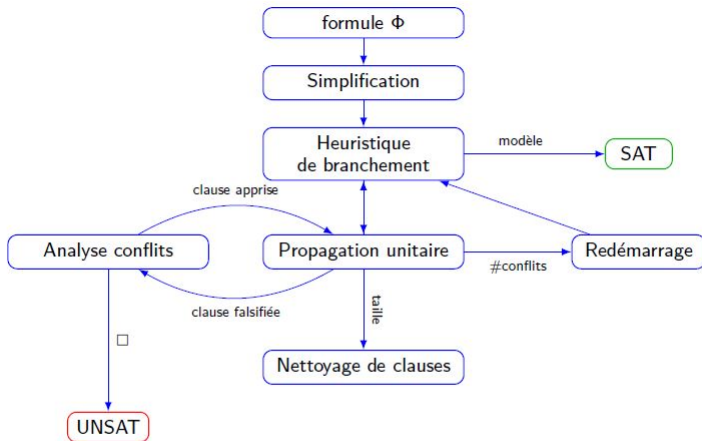
$$C_i = l_1 \vee \dots \vee l_n$$

- **Formule normale conjonctive (CNF)** : Une conjonction de clauses

$$\Phi = C_1 \wedge C_2 \wedge \dots \wedge C_n$$

- **Interprétation** : Une fonction qui associe une valeur à chaque variable.
- **Problème SAT** : Un problème qui permet de dire s'il existe une interprétation qui satisfait une formule CNF.

# Principe



# Principe

- Basé sur l'apprentissage de clauses

les clauses apprises  
ne peuvent être  
toutes conservées  
en mémoire



Suppression de  
certaines clauses  
selon des critères

# Apprentissage des clauses

- 3 principaux critères d'évaluation :
  - **Taille** : l'arité de la clause (le nombre de littéraux)
  - **Lbd** : (Literal Block Distance) : correspondant au nombre de niveaux de décision différents intervenant dans la génération de la clause apprise.
  - **Activité** : Mesure de l'apparition d'une clause dans les conflits

## Environnement de test

- **Système d'exploitation** : Ubuntu
- **Solveur** : Glucose 4.0  
<http://www.labri.fr/perso/lSimon/glucose/>
  - Politique de suppression de clauses agressive.
- **Timeout** : 15 minutes (900 secondes)
- **Instances testées** : 9 familles d'instances issues d'application réelle (2 instances de chaque famille)  
<https://baldur.iti.kit.edu/sat-competition-2016/>
  - Soit 18 instances testées



## Extraction des critères d'évaluation

- Au cours de la résolution, nous récupérons avant les réductions, pour chaque clause apprise :
  - sa taille
  - son activité
  - son lbd.
- Pour une instance, plusieurs réductions possible ; nous récupérons uniquement les informations avant :
  - La première réduction
  - La réduction au milieu de la recherche
  - La dernière réduction

## Corrélation entre critères

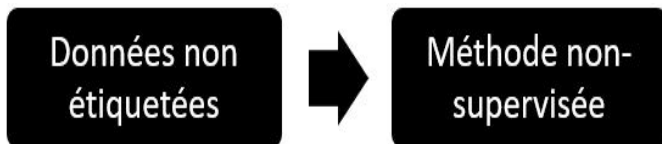
Instance	Réduction	lbd vs size	lbd vs activity	size vs activity
ablmulub16x4o	Début	0.82	-0.33	-0.32
	Milieu	0.59	-0.15	-0.11
	Fin	0.44	-0.15	-0.11
AProVE11-06	Début	0.73	-0.12	-0.10
	Milieu	0.71	-0.19	-0.17
	Fin	0.62	-0.14	-0.12
ak064modasbg2asisc	Début	0.88	-0.29	-0.29
	Milieu	0.79	-0.07	-0.04
	Fin	0.87	-0.07	-0.05
gaussian.c.75.smt2-cvc4	Début	0.68	0.00	0.06
	Milieu	0.71	0.00	-0.01
	Fin	0.69	0.00	-0.02
custmulsb2x32o	Début	0.81	-0.24	-0.24
	Milieu	0.77	-0.17	-0.17
	Fin	0.75	-0.09	-0.09
modgen-n200-m9	Début	0.88	-0.29	-0.22
	Milieu	0.86	-0.14	-0.14
	Fin	0.85	-0.07	-0.07

# Corrélation entre critères

Instance	Réduction	lbd vs size	lbd vs activity	size vs activity
ACG-20-5p1	Début	0.35	0.01	-0.08
	Milieu	0.27	-0.10	-0.06
	Fin	0.21	-0.10	-0.05
barman-pfile07-027	Début	0.38	-0.11	-0.14
	Milieu	0.23	-0.10	-0.11
	Fin	0.23	-0.10	-0.11
sokoban-p20.sas.cr.31	Début	0.40	-0.13	-0.13
	Milieu	0.29	-0.09	-0.10
	Fin	0.32	-0.06	-0.09

# Classification des clauses

- Pour la suppression des clauses, nous utilisons un classifieur sur des données non étiquetées.
- Le jeu de données est constitué comme tel :
  - **batch** : chaque clause apprise
  - **Features** : le lbd, la taille, l'activité



# Classification des clauses

- Clustering : Diviser les données en paquets homogènes
- Algorithme utilisé : k-means (k-moyennes)
- **1<sup>ère</sup> tentative** :
  - Classer les données en 2 classes :
    - Clauses qu'on peut supprimer
    - Celles qu'on doit garder.

## Classification des clauses

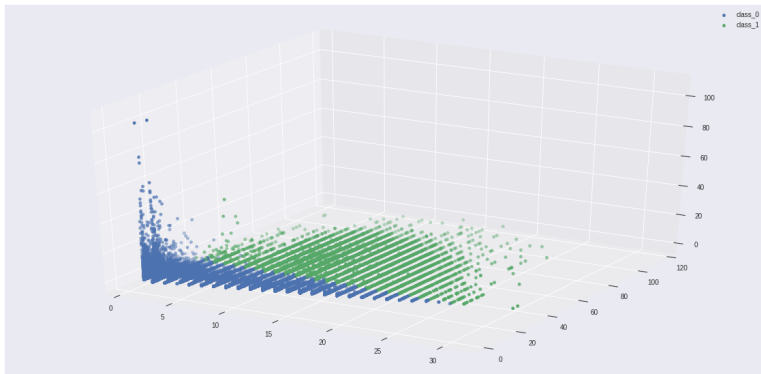


Figure 1 – Représentation de la classification en deux classes pour l'instance ablmulub16x4o

# Classification des clauses

- Nombre totale de clauses : 134014
- Coordonnées du centre de la 1<sup>ère</sup> classe : (6.03, 12.98, 0.34)
  - 111191 clauses
  - 82.97 % des données
- Coordonnées du centre de la 2<sup>ème</sup> classe : (12.55, 46.21, 0.04)
  - 22823 clauses
  - 17.03 % des données

## Classification des clauses

- Classification en 3 classes :

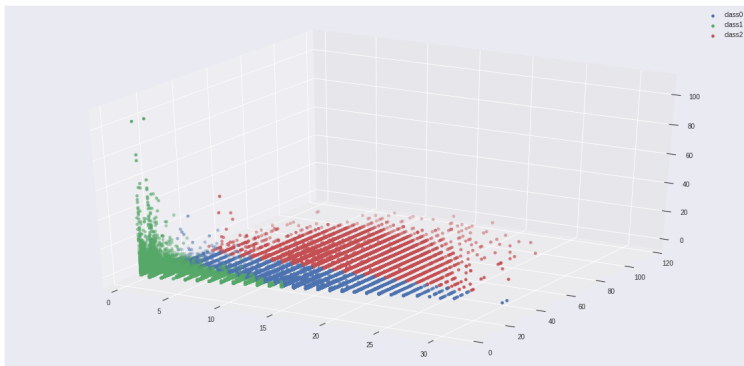


Figure 2 – Représentation de la classification en trois classes pour l'instance abmlub16x4o



# Classification des clauses

- Coordonnées du centre de la 1<sup>ère</sup> classe : (11.65, 22.90, 0.04)
  - 36805 clauses ( soit 27.46 %)
- Coordonnées du centre de la 2<sup>ème</sup> classe : (4.12, 10.35, 0.45)
  - 81813 clauses (soit 61.05 %)
- Coordonnées du centre de la 3<sup>ème</sup> classe : (12.42, 52.54, 0.04)
  - 15396 clauses (soit 11.49 %)

## Classification des clauses

- Classification en 4 classes :

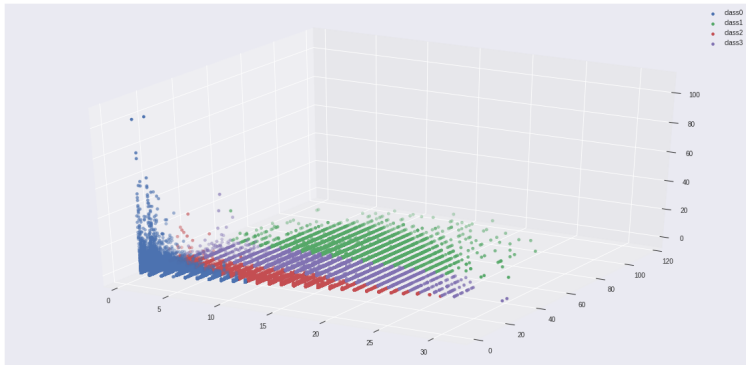


Figure 3 – Représentation de la classification en quatre classes pour l'instance ablmulub16x4o

## Classification des clauses

- Coordonnées du centre de la 1<sup>ère</sup> classe : (2.99, 9.59, 0.53)
  - 67246 clauses (50.18 %)
- Coordonnées du centre de la 2<sup>ème</sup> classe : [(12.20, 61.47, 0.03)]
  - 7951 clauses (5.93 %)
- Coordonnées du centre de la 3<sup>ème</sup> classe : (10.66, 17.77, 0.05)
  - 42220 clauses (31.5 %)
- Coordonnées du centre de la 4<sup>ème</sup> classe : (12.57, 37.08, 0.04)
  - 16597 clauses (12.38 %)

## Conclusion et perspective

- La division en plusieurs classes ne garantit pas la séparation selon un critère précis. Il est donc nécessaire d'effectuer une plus étude approfondie.
- Nous pouvons poursuivre une étude détaillée de la classification et l'intégrer à un solveur CDCL afin d'évaluer les performances.

Merci pour votre Attention !!!