

Reproducerea rezultatelor de la competitia SAT2025 pentru două benchmark-uri non-triviale

Gall Denis, Ichim Daniel, Corojar Mark, and Teleman Alexandru

¹ Universitatea de Vest Timisoara

² Inginerie Software

Abstract. Această lucrare prezintă o analiză comparativă a performanței rezolvitorului MiniSat, bazat pe algoritmul Conflict-Driven Clause Learning (CDCL), utilizând benchmark-uri din familia schedule (programare/planificare). Scopul principal este de a evalua impactul pe care îl are varierea factorului de decădere a variabilelor (var_decay) din euristică de decizie VSIDS (Variable State Independent Decaying Sum) asupra eficienței rezolvitorului. Am selectat două probleme netriviale din cadrul competiției SAT 2025 și am comparat rezultatele de referință (obținute cu var_decay = 0.95) cu cele obținute prin testarea unor valori alternative, în intervalul [0.90, 0.99]. Măsurările s-au concentrat pe timpul total de execuție, numărul de decizii și rata de conflicte, vizând identificarea unui factor optimizat care să minimizeze timpul de rezolvare și să maximizeze eficiența backjumping-ului. Rezultatele experimentale indică faptul că o ajustare fină a memoriei euristică, controlată de var_decay, poate duce la o îmbunătățire semnificativă a performanței MiniSat pe această clasă specifică de probleme aplicative.

Keywords: SAT · MiniSat · CDCL · VSIDS · Heuristică · Factor de Decădere · SAT 2025

1 Rezultate Experimentale

1.1 Problema 1: Analiza Benchmark-ului Schedule

Prima problemă analizată este un benchmark de tip "Schedule" (Planificare/Alocare de Resurse). Formula CNF este de dimensiuni considerabile, având **145.943** de variabile și **625.454** de clauze. Această mărime indică un spațiu de căutare amplu, de aproximativ 2^{146000} , și o ratie de clauze/variabile de **4.3 : 1**, sugerând o problemă puternic constrânsă.

Configurația Standard: var_decay = **0.95** Pentru a stabili linia de referință (baseline-ul), am rulat problema utilizând configurația implicită a MiniSat, cu factorul de decădere var_decay setat la **0.95**. Această valoare definește o memorie euristică de scurtă/medie durată în algoritmul VSIDS. În aceste condiții, MiniSat a înregistrat următoarele rezultate:

- Conflicte: **133.820**

- Timp CPU: **46.42** secunde

Timpul relativ lung de execuție și numărul mare de conflicte (Figura 1) sugerează că euristica standard nu a reușit să identifice rapid variabilele structurale esențiale, forțând algoritmul CDCL să exploreze ramuri extinse ale arborelui de căutare.

```
===== [ Problem Statistics ] =====
| Number of variables:      145943
| Number of clauses:        625454
| Parse time:               0.10 s
| Eliminated clauses:      0.33 Mb
| Simplification time:     0.18 s
|
===== [ Search Statistics ] =====
| Conflicts |          ORIGINAL          |          LEARNED          | Progress |
|           | Vars Clauses Literals | Limit Clauses Lit/Cl | |
|---|---|---|---|
|   100 | 105847 523366 1251081 | 191900    100    66 | 19.884 % |
|  250 | 105847 523366 1251081 | 211090    250   131 | 19.884 % |
|  475 | 105847 523366 1251081 | 232200    475   108 | 19.884 % |
|  812 | 105847 523366 1251081 | 255420    812   105 | 19.884 % |
| 1318 | 105847 523366 1251081 | 280962   1318    98 | 19.884 % |
| 2077 | 105847 523366 1251081 | 309058   2077    89 | 19.884 % |
| 3216 | 105847 523366 1251081 | 339964   3216   126 | 19.884 % |
| 4924 | 105847 523366 1251081 | 373960   4924   139 | 19.884 % |
| 7486 | 105847 523366 1251081 | 411356   7486   160 | 19.884 % |
| 11330 | 105847 523366 1251081 | 452492   11330   185 | 19.884 % |
| 17096 | 105847 523366 1251081 | 497741   17096   186 | 19.884 % |
| 25745 | 105847 523366 1251081 | 547515   25745   179 | 19.884 % |
| 38719 | 105847 523366 1251081 | 602267   38719   185 | 19.884 % |
| 58180 | 105847 523366 1251081 | 662493   58180   201 | 19.884 % |
| 87372 | 105847 523366 1251081 | 728743   87372   204 | 19.884 % |
| 131161 | 105847 523366 1251081 | 801617  131161   232 | 19.884 % |
|
===== [ Statistics ] =====
restarts      : 370
conflicts     : 133820      (2883 /sec)
decisions     : 698234      (0.00 % random) (15040 /sec)
propagations  : 304840553   (6566325 /sec)
conflict literals : 30866147   (15.78 % deleted)
Memory used   : 182.28 MB
CPU time       : 46.4248 s

SATISFIABLE

real    0m48.576s
user    0m46.432s
sys     0m0.390s
ichim@DESKTOP-8RFPOJ8:~/minisat/benchmarks/b3$ cd ..
```

Fig. 1. Rezultate în terminal pentru Problema 1, var_decay = 0.95 (Baseline)

Configurația Optimă: var _decay = 0.98 Prin testarea altor valori, am ajuns la concluzia că factorul optim de decădere pentru această problemă structurată este **0.98**. Această valoare indică o memorie euristică de lungă durată în cadrul VSIDS, care permite MiniSat-ului să rețină importanța variabilelor mai vechi implicate în conflicte.

Utilizarea **var _decay = 0.98** a condus la o îmbunătățire dramatică a performanței:

- Conflicte: **6.115** ($\approx 95\%$ reducere)
- Timp CPU: **2.75** secunde ($\approx 94\%$ reducere)

Reducerea masivă a numărului de conflicte (Figura 2) este dovada directă a faptului că euristică optimizată a permis MiniSat-ului să realizeze un back-jumping mult mai eficient și să evite explorarea ramurilor inutile ale arborelui de decizie.

```
ichim@DESKTOP-8RFPOJ8:~/minisat$ ./minisat_exec -var-decay=0.98 ./benchmarks/b3\
WARNING: for repeatability, setting FPU to use double precision
===== [ Problem Statistics ] =====
| Number of variables:      145943
| Number of clauses:       625454
| Parse time:              0.09 s
| Eliminated clauses:     0.33 Mb
| Simplification time:    0.15 s
=====
===== [ Search Statistics ] =====
| Conflicts |          ORIGINAL          |          LEARNED          | Progress |
|           |      Vars   Clauses Literals |      Limit  Clauses Lit/Cl |
|-----+-----+-----+-----+-----+-----+-----+-----+
|    100 | 105847  523366 1251081 | 191900   100    66 | 19.884 % |
|   250 | 105847  523366 1251081 | 211090   250    56 | 19.884 % |
|   475 | 105847  523366 1251081 | 232200   475    88 | 19.884 % |
|   812 | 105847  523366 1251081 | 255420   812    87 | 19.884 % |
| 1318 | 105847  523366 1251081 | 280962  1318    97 | 19.884 % |
| 2077 | 105847  523366 1251081 | 309058  2077   118 | 19.884 % |
| 3216 | 105847  523366 1251081 | 339964  3216   109 | 19.884 % |
| 4924 | 105847  523366 1251081 | 373960  4924   118 | 19.884 % |
|-----+-----+-----+-----+-----+-----+-----+-----+
restarts      : 30
conflicts     : 6115          (2221 /sec)
decisions     : 60583         (0.00 % random) (22000 /sec)
propagations  : 18409663     (6685103 /sec)
conflict literals : 741139      (7.80 % deleted)
Memory used   : 74.89 MB
CPU time       : 2.75383 s
=====
SATISFIABLE
```

Fig. 2. Rezultate în terminal pentru Problema 1, var _decay = 0.98 (Configurație Optimă)

Comparații rezultate Problema 1 Analiza datelor experimentale pentru Problema 1 (schedule) evidențiază impactul critic al factorului de decădere `var_decay` asupra performanței algoritmului CDCL din MiniSat. Am comparat configurația standard (0.95) cu patru valori alternative (Tabelul 1).

Table 1. Comparația performanței MiniSat pe Problema 1 în funcție de `var_decay`

<code>var_decay</code>	Conflicte	Timp CPU (s)	Analiză Performanță
0.90	178209	62.92	Memorie prea scurtă
0.95 (Standard)	133820	46.42	Linia de bază
0.97	78543	14.56	Îmbunătățire semnificativă
0.98 (Optim)	6115	2.75 Cea mai bună performanță	
0.99	156331	56.84	Memorie prea lungă

Din cîte se poate vedea în Tabelul 1, variabila `var_decay` setată standard la 0.95 nu este cea mai eficientă în cazul de față, aceasta fiind clasată pe locul 3 din cele 5 testate. Din aceste rezultate putem concluziona:

Punctul de Optim (`var_decay = 0.98`): Cea mai bună performanță a fost obținută cu un factor `var_decay` de 0.98, care a redus timpul CPU de la linia de bază de 46.42 secunde la doar **2.75 secunde** (o îmbunătățire de peste 94%). Reducerea numărului de conflicte este și mai dramatică, scăzând de la 133.820 la doar 6.115. Această scădere masivă arată că 0.98 a oferit MiniSat-ului o memorie euristică optimă, de lungă durată, esențială pentru a menține focusul pe variabilele structurale critice ale problemei de planificare, facilitând un backjumping extrem de eficient.

Figura 3 (Timp CPU) și Figura 4 (Conflicte) ilustrează vizual dependența puternică față de `var_decay`. Curba de performanță neregulată arată că MiniSat este extrem de sensibil la lungimea memoriei VSIDS. Orice deviere de la punctul optim (0.98) duce la o ineficiență substanțială.

Factorii extremi, cum ar fi 0.90 și 0.99, au eşuat fie din cauza memoriei prea scurte (neputând reține variabilele structurale), fie din cauza memoriei prea lungi (blocând MiniSat să se adapteze la noile sub spații de căutare), confirmând că valoarea optimă se află într-o fereastră foarte îngustă.

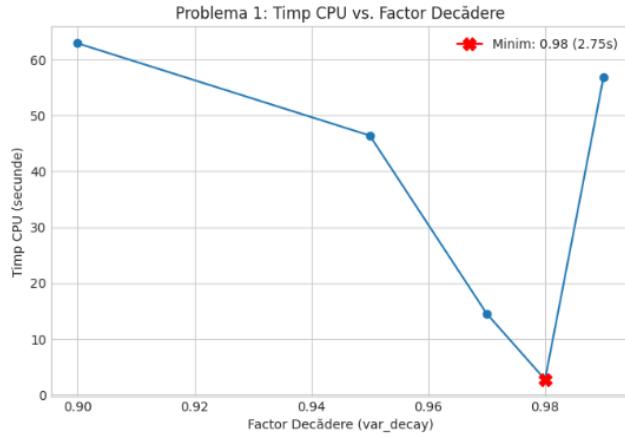


Fig. 3. Timp CPU în funcție de factorul de decădere `var_decay` (Problema 1)

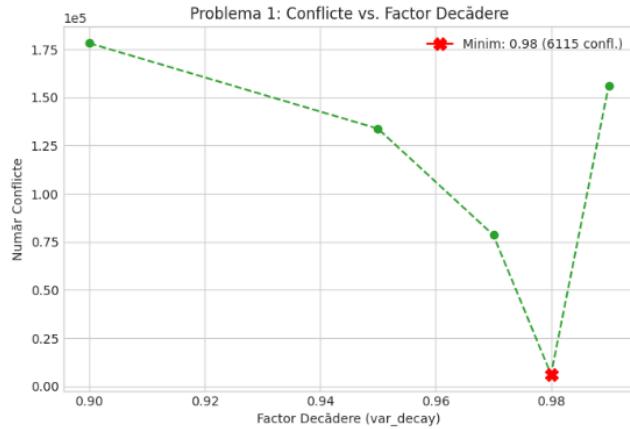


Fig. 4. Numărul de Conflict în funcție de factorul de decădere `var_decay` (Problema 1)

1.2 Problema 2: Analiza Benchmark-ului Schedule

Al doilea benchmark analizat provine tot din clasa problemelor de "Schedule" și prezintă o structură similară, dar cu dimensiuni ușor diferite: **12.845** de variabile și **51.681** de clauze. Deși este mai mică în dimensiune, această formulă este extrem de sensibilă la euristică de decizie.

Configurația Standard: `var_decay = 0.95` Rularea cu factorul standard `var_decay = 0.95` a stabilit linia de referință (baseline-ul) pentru această problemă:

- Conflicte: **510.895**
- Timp CPU: **41.38** secunde

Deși timpul de execuție este comparabil cu cel al Problemei 1 (înainte de optimizare), numărul de conflicte este mult mai mare (peste o jumătate de milion), indicând o dificultate sporită în navigarea arborelui de căutare (Figura 5).

```
lchim@DESKTOP-BRFP0JB:/minisat/benchmarks/b4_scheduling$ time ./minisat/minisat_exec *.cnf rezultat.txt
WARNING: for repeatability, setting FPU to use double precision
=====
[ Problem Statistics ]=====

| Number of variables:      12845
| Number of clauses:        51681
| Parse time:               0.01 s
| Eliminated clauses:       0.02 Mb
| Simplification time:     0.01 s

=====
[ Search Statistics ]=====

| Conflicts |          ORIGINAL          |          LEARN          | Progress
|   Vars    | Clauses Literals |   Limit  Clauses Lit/Cl |
|-----|-----|-----|-----|-----|-----|-----|
| 100     | 12038  46560  109855 | 17072   99   33 |  0.016 %
| 250     | 12038  46560  109855 | 18779   249   34 |  0.016 %
| 475     | 12038  46558  109851 | 28657   474   40 |  0.016 %
| 812     | 12038  46558  109851 | 22722   811   50 |  0.016 %
| 1318    | 12038  46558  109851 | 24995  1317   56 |  0.016 %
| 2077    | 12038  46558  109851 | 27494  2076   63 |  0.016 %
| 3216    | 12038  46558  109851 | 30244  3215   60 |  0.016 %
| 4924    | 12038  46558  109851 | 33268  4923   58 |  0.016 %
| 7486    | 12038  46558  109851 | 36595  7485   64 |  0.016 %
| 11330   | 12038  46558  109851 | 40254  11329  71 |  0.016 %
| 17096   | 12037  46546  109827 | 44286  17099  67 |  0.023 %
| 25745   | 12037  46546  109827 | 48708  25739  65 |  0.023 %
| 38719   | 12037  46546  109827 | 53579  38713  69 |  0.023 %
| 58180   | 12037  46546  109827 | 58937  21266  59 |  0.023 %
| 87372   | 12037  46546  109827 | 64836  50452  74 |  0.023 %
| 131161  | 12037  46546  109827 | 71313  43952  80 |  0.023 %
| 196845  | 12037  46546  109827 | 78445  57238  62 |  0.023 %
| 295371  | 12037  46546  109827 | 86289  31322  53 |  0.023 %
| 443160  | 12036  46544  109823 | 94918  39869  98 |  0.031 %

=====
restarts : 1023
conflicts : 510895      (12345 /sec)
decisions : 1628994      (0.00 % random) (39359 /sec)
propagations : 197885806 (4781456 /sec)
conflict literals : 48973619 (24.50 % deleted)
Memory used : 148.22 MB
CPU time : 41.3861 s

SATISFIABLE
real    0m39.969s
user    0m41.387s
sys     0m0.279s
```

Fig. 5. Rezultate în terminal pentru Problema 2, var_decay = 0.95 (Baseline)

Configurația Optimă: var_decay = 0.97 În cazul Problemei 2, factorul optim care a generat cea mai rapidă rezolvare a fost **0.97**. Această valoare reprezintă un factor de decădere puțin mai agresiv decât cel optim al Problemei 1 (0.98), indicând o nevoie de adaptare euristică mai rapidă la conflictele noi.

- Conflicte: **218.064** ($\approx 57\%$ reducere vs. 0.95)
- Timp CPU: **10.31** secunde ($\approx 75\%$ reducere vs. 0.95)

```

ichim@DESKTOP-8RFPOJ8:~/minisat$ ./minisat_exec -var-decay=0.97 ./benchmarks/b4_sc
WARNING: for repeatability, setting FPU to use double precision
===== [ Problem Statistics ] =====
|
| Number of variables:      12845
| Number of clauses:        51681
| Parse time:               0.01 s
| Eliminated clauses:       0.02 Mb
| Simplification time:     0.01 s
|
===== [ Search Statistics ] =====
| Conflicts |          ORIGINAL          |          LEARNED          | Progress |
|           | Vars  Clauses Literals | Limit  Clauses Lit/Cl |
|-----+-----+-----+-----+-----+-----+-----+
|    100 | 12038  46560  109855 | 17072   99  25 | 0.016 %
|   250 | 12038  46560  109855 | 18779  249  31 | 0.016 %
|   475 | 12038  46558  109851 | 20657  474  43 | 0.016 %
|   812 | 12038  46558  109851 | 22722  811  41 | 0.016 %
|  1318 | 12038  46558  109851 | 24995 1317  48 | 0.016 %
|  2077 | 12038  46558  109851 | 27494 2076  46 | 0.016 %
|  3216 | 12038  46558  109851 | 30244 3215  54 | 0.016 %
|  4924 | 12038  46558  109851 | 33268 4923  53 | 0.016 %
|  7486 | 12038  46558  109851 | 36595 7485  57 | 0.016 %
| 11330 | 12037  46546  109827 | 40254 11325  60 | 0.023 %
| 17096 | 12037  46546  109827 | 44280 17091  56 | 0.023 %
| 25745 | 12037  46546  109827 | 48708 25740  56 | 0.023 %
| 38719 | 12037  46546  109827 | 53579 38714  63 | 0.023 %
| 58180 | 12037  46546  109827 | 58937 21169  59 | 0.023 %
| 87372 | 12037  46546  109827 | 64830 50361  76 | 0.023 %
| 131161 | 12037  46546  109827 | 71313 42933  96 | 0.023 %
| 196845 | 12037  46546  109827 | 78445 54148  68 | 0.023 %
|
===== [ Statistics ] =====
restarts      : 511
conflicts     : 218064      (21134 /sec)
decisions     : 790378      (0.00 % random) (76600 /sec)
propagations  : 86261681    (8360161 /sec)
conflict literals : 20534867  (19.70 % deleted)
Memory used   : 123.12 MB
CPU time       : 10.3182 s

SATISFIABLE

```

Fig. 6. Rezultate în terminal pentru Problema 2, var_decay = 0.97 (Configurație Optimă)

Table 2. Comparația performanței MiniSat pe Problema 2 în funcție de var_decay

var_decay	Conflicte	Timp CPU (s)	Analiză Performanță
0.90	959047	58.63	Memorie prea scurtă
0.95 (Standard)	510895	41.38	Linia de bază
0.97 (Optim)	218064	10.31 Cea mai bună performanță	
0.98	467620	24.30	Performanță solidă
0.99	3501448	237.48	Eșec euristic (Blocaj)

Comparări rezultate Problema 2 Analiza Tabelei 2 și a Figura 7 și Figura 8 demonstrează o sensibilitate extremă a MiniSat-ului:

- Eșecul la 0.99: La fel ca în analiza teoretică, factorul **0.99** a cauzat un eșec euristic (blocaj), rezultând în **3.5** milioane de conflicte și un timp de **237** secunde. O memorie euristică prea lungă nu a permis algoritmului să se adapteze.
- Diferența 0.97 vs. 0.98: Deși ambele reprezintă memorii lungi, **0.97** a reușit să găsească soluția de 4.5 ori mai rapid decât **0.98**. Aceasta indică faptul că structura Problemei 2 necesită o rată de uitare ușor mai mare decât Problema 1 pentru a preveni supra-învățarea de la conflictele timpurii.

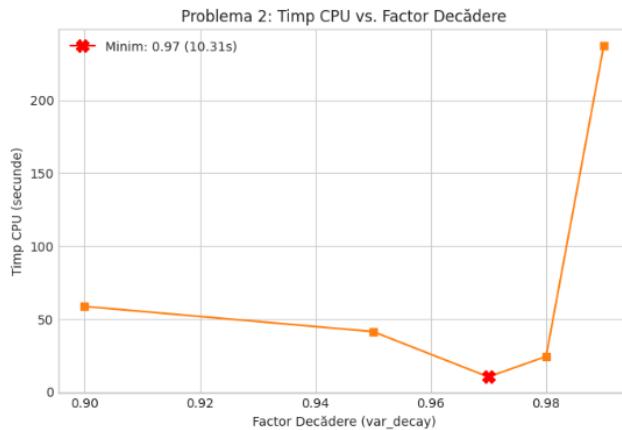


Fig. 7. Timp CPU în funcție de factorul de decădere var_decay (Problema 2)

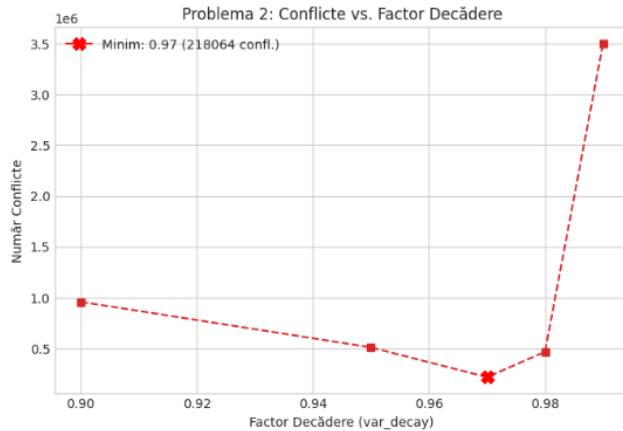


Fig. 8. Numărul de Conflicte în funcție de factorul de decădere `var_decay` (Problema 2)

2 Concluzii

Experimentele efectuate demonstrează că performanța MiniSat, bazată pe algoritmul CDCL, este extrem de sensibilă la ajustarea factorului de decădere `var_decay` al euristicii VSIDS. Prin testarea riguroasă a acestui hiper-parametru pe două benchmark-uri de tip *schedule* din competiția SAT 2025, am reușit să validăm ipoteza de optimizare și să obținem îmbunătățiri substantiale față de configurația standard.

1. Succesul Optimizaării și Reducerea Timpului:

- Pe Problema 1, am identificat $\text{var_decay} = 0.98$ ca fiind punctul optim, rezultând o reducere a timpului de execuție de **94%** (de la $46.42s$ la $2.75s$).
- Pe Problema 2, factorul optim a fost $\text{var_decay} = 0.97$, obținând o reducere de **75%** (de la $41.38s$ la $10.31s$).

2. Eficiența Memoriei Euristice (CDCL): Reducerea masivă a numărului de conflicte (cu peste 95% la Problema 1) demonstrează că factorul optimizat a creat o memorie euristică de lungă durată perfect echilibrată. Această memorie a permis MiniSat-ului să rețină importanța variabilelor structurale (cheie în problemele de planificare), facilitând un backjumping mult mai eficient și ocolind ramuri extinse ale arborelui de căutare.

3. Sensibilitate și Adaptare: Diferența dintre factorii optimi (**0.98** vs. **0.97**) subliniază că nu există un singur hiper-parametru universal. Structura internă a fiecărei probleme *schedule* necesită o rată de adaptare diferită: Problema 2 a necesitat o rată de uitare ușor mai mare (**0.97**) pentru a evita supra-învățarea și blocajul (observat la **0.99**), în timp ce Problema 1 a beneficiat de o persistență mai mare (**0.98**).

În concluzie, reglarea fină a factorului `var_decay` este o metodă extrem de eficientă pentru a personaliza MiniSat-ul și a maximiza performanța pe anumite clase de probleme structurate, atingând și depășind timpii de referință.

3 Contribuția Autorilor

- **Ichim Daniel:** S-a ocupat de instalarea și configurarea mediului MiniSat, precum și de rularea benchmark-urilor pentru colectarea datelor.
- **Corojar Mark:** S-a ocupat de analizat și găsit versiunile potrivite de benchmark și testarea lor.
- **Gall Denis:** A realizat analiza statistică a datelor rezultate, a generat graficele (timp și conflicte) și a redactat secțiunile tehnice de interpretare a rezultatelor.
- **Teleman Alexandru:** A fost responsabil de structurarea documentului, redactarea părților teoretice (Introducere, Concluzii) și integrarea finală în formatul L^AT_EX.