Evaluarea MiniSat: Performanță și Provocări în Rezolvarea Problemelor SAT

Grosu Damian[1], Scoropad Iulian[2], Daniel Dobrescu[3], and Anton Constantin-Adrian[4]

[1] Universitatea de Vest Timisoara, Bulevardul Vasile Pârvan 4, Timișoara 300223
costi.dobrescu01@e-uvt.ro,
[2] iulian.scoropad00@e-uvt.ro,
[3] damian.grosu01@e-uvt.ro
http://www.uvt.ro

Abstract. Raportul dat se concentrează pe analiza problemei satisfiabilității booleene (SAT), prima problemă recunoscută ca NP-completă, și pe aplicabilitatea acesteia în diverse domenii. Documentul prezintă în detaliu funcționarea și îmbunătățirile aduse MiniSat, un solver SAT minimalist și eficient,dezvoltat de Niklas Eén și Niklas Sörensson. Raportul include o descriere a structurilor și algoritmilor MiniSat, ghiduri pentru instalare, precum și metodologia utilizată pentru testarea performantelor acestuia. Deasemenea MiniSat va fi testat pe un set de benchmark-uri din competiția SAT 2024, analizând performanța MiniSat în termeni de satisfiabilitate, relevând constrângerile de timp, hardware și discutând eficiența acestuia în diferite scenarii. Concluziile care le vom sublinia în urma realizării proiectului vor arăta atât succesul în rezolvarea problemelor simple, dar și dificultățile întâlnite în cazuri complexe, oferind perspective asupra îmbunătățirilor viitoare.

1 Introducere

Rezolvarea problemei de satisfiabilitate booleeană (SAT) reprezintă un punct central în cercetarea în domeniul informaticii, fiind primul exemplu de problemă NP-completă. Termenul NP vine de la nondeterministic polynomial time (timp polinomial nedeterminist) iar în teoria complexității calculului, NP este o clasă de probleme decizionale care are următoarele caracteristici:

Verificare eficientă: Dacă cineva îți oferă o soluție (un „certificat") pentru o problemă din NP, poți verifica dacă soluția este corectă în timp polinomial folosind un algoritm determinist. Cu alte cuvinte, soluția poate fi verificată rapid (eficient) de un calculator.

Verificare eficientă: Dacă cineva îți oferă o soluție (un „certificat") pentru o problemă din NP, poți verifica dacă soluția este corectă în timp polinomial folosind un algoritm determinist. Cu alte cuvinte, soluția poate fi verificată rapid (eficient) de un calculator.

SAT (Satisfiability Problem) este o problemă din logica propozițională, care implică determinarea existenței unei asignări de variabile booleene astfel încât o formulă dată să fie adevărată. Această problemă a fost definită formal ca fiind NP-completă de către Cook în 1971. De atunci, SAT a devenit un subiect de mare interes datorită implicațiilor teoretice și practice.

SAT Solvers sunt folosiți pe scară largă într-o varietate de domenii, cum ar fi:

Verificarea formală a hardware-ului și software-ului (Model Checking)
SAT Solvers sunt utilizați pentru a verifica dacă un sistem hardware sau software respectă specificațiile dorite, identificând eventualele erori.
Exemplu: În verificarea microprocesoarelor, un SAT Solver poate fi utilizat pentru a verifica dacă circuitele logice implementate funcționează conform designului specificat, detectând potențiale situații de blocaj sau comportamente incorecte.
Aplicație concretă: Intel folosește SAT Solvers pentru a valida designul procesoarelor înainte de producția de masă.
În software, SAT Solvers ajută la analiza și detectarea bug-urilor în programe critice, cum ar fi cele pentru aviație sau medicină

Planificare automată și inteligență artificială
SAT Solvers sunt folosiți pentru a rezolva probleme complexe de planificare, în care trebuie găsită o secvență optimă de acțiuni pentru a atinge un obiectiv.
Exemplu: Un robot care planifică rutele de curățare într-o clădire trebuie să determine traseele eficiente care să acopere toate camerele, evitând obstacolele. SAT Solvers pot rezolva astfel de probleme modelând restricțiile și obiectivele ca o formulă logică.
Aplicație concretă: În jocurile video, SAT Solvers pot fi utilizați pentru a genera strategii optime pentru personaje controlate de inteligența artificială.

Optimizarea și rezolvarea problemelor de satisfacție a constrângerilor (CSP)

SAT Solvers sunt esențiali în problemele de optimizare care implică multe restricții. Acestea pot varia de la alocarea resurselor până la programarea unor activități.
Exemplu: Planificarea orarului școlar într-o școală mare, astfel încât să se evite suprapunerea profesorilor, sălilor sau grupurilor de studenți. SAT Solver-ul poate rezolva problema transformând constrângerile într-o formulă logică satisfiabilă.
Aplicație concretă: Companii aeriene folosesc SAT Solvers pentru optimizarea rutelor de zbor și a programelor echipajelor, reducând costurile operaționale.

Criptografie și securitate cibernetică
Exemplu: Testarea algoritmilor de criptare pentru a descoperi chei slabe. SAT Solvers pot încerca să determine dacă o anumită cheie secretă poate fi derivată mai ușor decât ar trebui în mod normal.
Aplicație concretă: În analiza malware, SAT Solvers pot fi folosiți pentru a identifica dacă un set de reguli de securitate este satisfăcut de un comportament suspect observat într-un sistem.

2 Descrierea problemei

Problema principală abordată de un SAT Solver este determinarea satisfiabilității unei formule logice. Satisfiabilitatea presupune existența unei configurații de valori de adevăr (True/False) pentru variabilele formulei care să conducă la un rezultat adevărat. Dacă o astfel de configurație există, formula este considerată satisfiabilă; în caz contrar, este nesatisfiabilă.

Rezolvarea problemelor de satisfiabilitate este importantă deoarece multe alte probleme complexe pot fi reduse la SAT. În cazul nostru, ne propunem să verificăm dacă o formulă logică dată este satisfiabilă sau nu. Formula logică este exprimată în formă normală conjunctivă, compusă din mai multe clauze care sunt, la rândul lor, disjuncții de literali.

3 Minisat

3.1 Introducere în Minisat

MiniSat este un solver minimalist și open-source destinat rezolvării problemelor de satisfiabilitate logică (SAT - Satisfiability). Dezvoltat inițial de Niklas Eén și Niklas Sörensson, MiniSat a devenit rapid un punct de referință în domeniul solverelor SAT datorită simplității, performanței și modularității sale. A obținut realizări semnificative în competițiile internaționale de SAT, cum ar fi SAT Competition, unde a avut multiple performanțe, inclusiv câștigarea competiției din 2005. Aceste succese au demonstrat că MiniSat este un solver eficient, capabil să abordeze o gamă largă de instanțe de satisfiabilitate.SAT-ul este o problemă de bază în calculul și logica matematică, constând în determinarea dacă o formulă booleană poate fi evaluată ca adevărată (satisfiabilă) printr-o atribuire adecvată a valorilor de adevăr variabilelor sale. Popularitatea MiniSat derivă din implementarea eficientă a tehnicilor moderne SAT și din capacitatea sa de a rezolva rapid probleme complexe.

MiniSat este aplicabil în numeroase domenii datorită capacității sale de a analiza expresii logice complexe și de a determina satisfiabilitatea acestora. Printre cele mai semnificative aplicații ale MiniSat se numără:

În ingineria software și hardware, MiniSat este utilizat pentru verificarea formală a modelelor logice și a circuitelor digitale. Prin transformarea problemelor în formă SAT, proiectanții pot verifica proprietăți precum coerența, corectitudinea și consistența modelelor.

În optimizarea planificării și alocării resurselor în proiecte complexe, cum ar fi gestionarea orarelor, planificarea în producție sau alocarea resurselor în centre de date.

În domeniul criptografiei, MiniSat este folosit pentru analizarea și verificarea algoritmilor criptografici. Poate detecta vulnerabilități în sisteme prin formularea problemelor de atac sau rezistență sub formă de instanțe SAT.

În bioinformatică, problemele de asamblare a genomului sau predicția structurilor moleculare pot fi exprimate în format SAT, iar MiniSat oferă o platformă eficientă pentru rezolvarea acestor probleme.

În inteligență artificială, MiniSat este folosit pentru raționamente logice și rezolvarea automată a problemelor de satisfiabilitate, având aplicații în cercetarea sistemelor de planificare automată, generarea de soluții optime în probleme de optimizare și analiza complexității problemelor combinatorii.

3.2 Algoritmi utilizați în MiniSAT

MiniSat își bazează performanța pe doi algoritmi principali: CDCL (Conflict-Driven Clause Learning) și DPLL (Davis-Putnam-Logemann-Loveland).

CDCL extinde DPLL prin învățarea din conflicte, reducând spațiul de căutare prin generarea de clauze noi care elimină cauzele conflictelor. În timpul căutării, propagarea unitară este optimizată prin tehnici precum Blocking Literals, iar variabilele sunt alese folosind euristici precum VSIDS, care prioritizează cele implicate în conflicte recente.

DPLL rămâne fundamentul căutării, utilizând propagarea unitară și backtracking pentru explorarea sistematică a soluțiilor, însă CDCL adaugă mecanisme precum sărituri peste niveluri multiple în cazul conflictelor și restarturi adaptive pentru a preveni stagnarea. În plus, MiniSat permite rezolvarea incrementală a problemelor, economisind resurse în aplicațiile dinamice. Astfel, integrarea acestor tehnici face procesul mai eficient și mai robust.

3.3 MiniSat și îmbunătățiri aduse în versiunea 2.2.0

În versiunea 2.2.0 a MiniSat, au fost implementate o serie de îmbunătățiri semnificative, menite să îmbunătățească performanța, portabilitatea și gestionarea resurselor. Printre cele mai importante modificări se numără:

3.4 Analiză Detaliată a Structurii și Modului de Lucru al MiniSat

Codul din programul MiniSat ruleaza secvențial pentru a rezolva problemele logice de satisfiabilitate. Programul începe prin configurarea solver-ului și citirea

fișierului de intrare care conține formula logică. În continuare se deduc valorile variabilelor prin propagarea unitară și se elimină clauzele satisfăcute pentru a reduce complexitatea căutării. Apoi se aleg variabilele pentru a explora soluțiile posibile, iar dacă apar conflicte, solver-ul revine asupra deciziilor anterioare. Când un conflict este găsit, se analizează cauza care a produs acel conflict și se adaugă o nouă clauză pentru a preveni erorile viitoare. Salvarea clauzelor ajută la optimizarea căutării, iar în cazul unor conflicte repetate, solver-ul revine la nivele anterioare pentru a încerca soluții noi. Reducerea informațiilor despre clauze economisește memorie, iar restarturile sunt folosite pentru a evita blocarea solver-ului. În cazul în care nu apar conflicte, se verifică satisfiabilitatea formulei, stabilind astfel dacă soluția găsită este validă sau nu.

În continuare vom prezenta detaliat pasii care se executa în programul Minisat în timpul rulării unui fișier :

1. **Inițializarea și setarea opțiunilor**
   Codul începe cu funcția main, care inițializează solver-ul și configurează opțiunile necesare pentru rularea acestuia. În main, se folosesc funcțiile setUsageHelp și parseOptions pentru a defini și interpreta opțiunile de utilizare. Aceste opțiuni includ gradul de detaliere al mesajelor (verbosity), limitele de timp CPU (cpu_lim) și memorie (mem_lim), și validarea antetului DIMACS (strictp). Configurarea acestor opțiuni controlează comportamentul solver-ului în timpul execuției.
2. **Citirea și parsarea fișierului de intrare**
   Solver-ul citește fișierul de intrare care conține expresia logică de verificat. Acest fișier poate fi în format plain text sau gzipped DIMACS. Funcția parse_DIMACS este responsabilă pentru parsarea fișierului și extragerea clauzelor SAT. Acest proces este foarte important pentru inițializarea clauzelor care vor fi utilizate în procesul de rezolvare.
3. **Inițializarea solver-ului**
   După parsarea fișierului de intrare, solver-ul este inițializat. Obiectul Solver este creat și configurat cu opțiunile setate anterior. Solver-ul este pregătit pentru a începe procesul de rezolvare a problemei SAT, iar obiectul Solver se ocupă de toate operațiunile ulterioare.
4. **Propagarea unitară și simplificarea**
   Primul pas în procesul de rezolvare este propagarea unitară, implementată de funcția propagate(). Scopul acestei funcții este de a deduce noi valori ale variabilelor bazate pe clauzele actuale. Funcția verifică fiecare clauză pentru a vedea dacă literalul negat determină clauza să fie satisfăcută sau nu. Dacă niciunul dintre ceilalți literați din clauză nu este adevărat, literalul rămas trebuie să fie adevărat pentru a evita un conflict. Solver-ul aplică și simplificări pentru a reduce spațiul de căutare. Funcția simplify() elimină clauzele satisfăcute și face alte optimizări pentru a accelera procesul de rezolvare. Aceste simplificări sunt esențiale pentru menținerea eficienței solver-ului.
5. **Procesul de decizie și crearea nivelelor de decizie** Pentru a explora diferite asignări ale variabilelor, solver-ul utilizează un proces iterativ de decizie. Funcția pickBranchLit() selectează următoarea variabilă care va fi

utilizată pentru decizie, bazată pe diverse euristici. După selectarea variabilei, solver-ul creează un nou nivel de decizie prin funcția newDecisionLevel(). Fiecare nivel de decizie reprezintă un moment în care o variabilă este atribuită fără propagare imediată, menținând astfel o structură ierarhică a deciziilor care facilitează backtracking-ul.

6. **Detectarea și analiza conflictelor** În timpul procesului de decizie și propagare, solver-ul poate întâlni conflicte, adică situații în care o clauză devine complet falsă. Când se întâmplă acest lucru, funcția analyze() examinează conflictul și generează o clauză de conflict care explică de ce decizia respectivă a condus la conflict. Această analiză implică examinarea tuturor clauzelor implicate și crearea unei clauze noi care explică conflictul.

7. **Memorarea clauzelor** Clauza de conflict rezultată din analiza conflictului. Aceasta este salvată în memoria calculatorului și utilizată de solver pentru a preveni conflicte similare în viitor. Această tehnică, cunoscută sub numele de învățarea clauzelor, este esențială pentru îmbunătățirea eficienței solver-ului. Funcțiile varBumpActivity() și claBumpActivity() ajustează activitatea variabilelor și clauzelor, reflectând importanța lor în rezolvarea problemei. De exemplu, variabilele și clauzele implicate frecvent în conflicte primesc o activitate mai mare, ceea ce influențează selecțiile viitoare ale solver-ului.

8. **Backtracking**
   În cazul unui conflict, solver-ul utilizează backtracking non-cronologic pentru a reveni la un nivel anterior de decizie. Funcția cancelUntil() este responsabilă pentru resetarea stării solver-ului până la nivelul de decizie specificat. Aceasta permite solver-ului să evite deciziile care au condus la conflict și să exploreze alte posibilități. Pentru a menține eficiența, solver-ul aplică și tehnici de reducere a clauzelor memorate. Funcția reduceDB() sortează clauzele învățate și le elimină pe cele mai puțin active, menținând astfel clauzele importante și reducând consumul de memorie.

9. **Gestionarea restarturilor**
   Restarturile sunt o tehnică utilizată pentru a evita blocarea solver-ului în anumite regiuni ale spațiului de căutare. Funcția luby() implementează secvența Luby pentru a determina intervalele de restart, echilibrând frecvența acestora pentru a îmbunătăți eficiența generală a solver-ului. Funcția solve_() controlează procesul de rezolvare și gestionează restarturile periodice, asigurându-se că solver-ul explorează eficient spațiul de căutare.

10. **Garbage Collector-ul folosit în MiniSat**
    Pentru a gestiona eficient memoria, solver-ul implementează tehnici de colectare a datelor reziduale. Funcția checkGarbage() verifică dacă există memorie inutilizată care trebuie colectată, iar funcția garbageCollect() realizează colectarea efectivă prin eliberarea și realocarea memoriei ocupate de clauze. Aceste funcții asigură utilizarea eficientă a resurselor de memorie, prevenind scăderea performanței din cauza acumulării de clauze inutile.

3.5 Instalarea Minisat on PC

Linux Pe Linux, MiniSat poate fi instalat rapid folosind compilarea codului sursă. Pentru realizarea acestui lucru trebuie sa urmarim cațiva pași:

1. Instalează compilatorul și uneltele de build:
```
sudo apt install g++ make
```
2. Clonează repository-ul MiniSat:
```
cd core make
```
3. Compilează proiectul:
```
./minisat
```

Mac Pe MAC, poate fi instalat usor utilizând Homebrew. Homebrew este un manager de pachete pentru macOS (și Linux) care simplifică instalarea softwareului și a uneltelor de dezvoltare direct din linia de comandă. Pentru configurarea MiniSAT va trebui sa urmarim câțiva pași:

1. Asigură-te că Homebrew este instalat. Dacă nu, instalează-l rulând comanda:
```
/bin/bash -c "$(curl -fsSL https://raw.
githubusercontent.com/Homebrew/install/HEAD/
install.sh)"
```
2. Instalează MiniSat folosind Homebrew:
```
brew install minisat
```

Windows MiniSat nu are o versiune nativă pentru Windows, dar poate fi rulat folosind soluții precum WSL (Windows Subsystem for Linux) sau MinGW. Pentru instalarea MiniSAT utilizand WSL vom parcurge următorii pași:

1. Activează WSL rulând comanda:
```
wsl --install
```
2. Instalează o distribuție Linux, precum Ubuntu, din Microsoft Store.
3. După instalare, deschide terminalul Linux și urmează pașii pentru Linux.

4 Algoritmii DPLL și CDCL

Algoritmii Davis-Putnam-Logemann-Loveland (DPLL) și Conflict-Driven Clause Learning (CDCL) sunt metode esențiale pentru determinarea satisfiabilității formulelor în formă normală conjunctivă (CNF). Acestea stau la baza majorității solverelor SAT moderne, având aplicații diverse în verificarea modelelor, bioinformatică și criptografie. În continuare vom analiza structura, componentele și modul de funcționare al celor doi algoritmi, ilustrând implementarea acestora în cadrul codului MiniSat, în special în fișierul Solver.cc unde la rândul său sunt dezvoltați algoritmii.

4.1 Algoritmul DPLL

DPLL este un algoritm de căutare în adâncime care combină tehnici de deducere și decizie pentru a găsi o evaluare satisfăcătoare a variabilelor din formule CNF. În cadrul acestui proces, DPLL parcurge sistematic spațiul de căutare, utilizând propagarea literaliilor unitari, analizând conflictele, și făcând backtracking pentru a explora soluțiile posibile. Principalele componente ale algoritmului DPLL sunt:

1. **Propagarea unității:** Atunci când o clauză conține un singur literal neasignat, valoarea acestuia este stabilită astfel încât să satisfacă clauza respectivă. Această operațiune poate declanșa alte asignări implicite în lanț, ceea ce poate duce fie la un conflict, fie la o soluție satisfăcătoare. Funcția propagate din cod este responsabilă pentru acest proces. Aceasta parcurge literalii, aplică propagarea și returnează clauza conflictuală dacă se detectează un conflict.
2. **Heuristica de decizie:** Algoritmul selectează variabile pentru decizie, alegând în mod euristic sau aleatoriu variabila care urmează să fie asignată, pentru a accelera căutarea. Funcția pickBranchLit din codul MiniSat implementează această etapă, având în vedere parametrii random_var_freq și euristica de activitate a variabilelor.
3. **Backtracking-ul:** Dacă algoritmul întâlnește un conflict, acesta revine la un nivel anterior de decizie pentru a explora alte opțiuni posibile. Funcția cancelUntil realizează acest backtracking în cod, eliminând atribuțiile variabilelor până la un anumit nivel, astfel încât căutarea să fie reluată de la o stare anterioară.
4. **Corectitudinea algoritmului:** DPLL este garantat să se termine fie cu o evaluare satisfăcătoare a variabilelor, fie cu verdictul că formula este nesatisfiabilă. Corectitudinea sa se bazează pe faptul că algoritmul explorează complet spațiul de soluții posibil și că orice clauze adăugate sunt echivalente cu formula inițială.

DPLL reprezintă o metodă fundamentală care, deși simplă, constituie baza pentru majoritatea solverelor SAT și este punctul de plecare pentru metode mai complexe, precum CDCL.

4.2 Algoritmul CDCL

Conflict-Driven Clause Learning (CDCL) este o extensie avansată a DPLL, care include optimizări semnificative pentru a evita conflictele și a gestiona eficient spațiul de căutare. CDCL utilizează învățarea clauzelor conflictuale, backtracking-ul non-cronologic și tehnici de restart pentru a îmbunătăți substanțial eficiența rezolvării problemelor SAT.

Principalele componente ale algoritmului CDCL sunt:

1. **Învățarea clauzelor din conflicte:** Atunci când apare un conflict, CDCL analizează cauza și generează o clauză nouă care reflectă condițiile conflictuale. Aceasta este stocată ca „clauză învățată” și va împiedica re-apariția

aceluiași conflict în viitor. Funcția analyze din cod implementează acest mecanism de învățare, derivând clauza conflictuală din condițiile care au dus la conflict.
2. **Propagarea unității și analiza conflictelor:** Similar DPLL, CDCL folosește propagarea literaliilor unitari pentru a asigna variabile și pentru a identifica conflictele. Însă, în cazul unui conflict, CDCL utilizează grafuri de implicații pentru a determina „punctul de implicație unică" (UIP), care reprezintă condiția specifică ce a generat conflictul. Aceasta permite un backtracking mai eficient, având ca rezultat reducerea numărului de conflicte repetitive, în cod funcția litRedundant asigură optimizarea clauzelor învățate, eliminând literalii redundanți care nu contribuie direct la conflictul curent.
3. **Heuristici de decizie și ștergerea clauzelor:** CDCL utilizează euristici avansate, cum ar fi VSIDS (Variable State Independent Decaying Sum), care prioritează variabilele recente în conflicte. De asemenea, funcția reduceDB elimină clauzele învățate mai puțin utile, pentru a optimiza memoria utilizată și pentru a menține eficiența solver-ului.
4. **Restarturi aleatorii:** În anumite instanțe de SAT dificile, CDCL poate efectua restarturi periodice pentru a evita blocajele în sub-arborii de căutare. Acest lucru este realizat prin funcția luby, care definește intervalele de timp pentru restarturi. Restarturile permit solver-ului să exploreze mai eficient diverse regiuni ale spațiului de căutare.

Aceste funcții, combinate cu strategii avansate de decizie și propagare, permit CDCL să rezolve probleme SAT complexe cu eficiență sporită, în special acolo unde algoritmul DPLL deja nu mai este eficient.

5 Benchmark SAT Competition 2024

Competiția SAT 2024 reprezintă un eveniment anual de referință în domeniul rezolvării problemelor de satisfiabilitate booleană (SAT). Aceasta este organizată ca un eveniment satelit al Conferinței SAT 2024 și continuă tradiția competițiilor SAT și provocărilor SAT-Race.

În ultimii ani, domeniul SAT a înregistrat progrese remarcabile, transformând probleme considerate anterior imposibil de rezolvat în cazuri rezolvabile de rutină. Aceste avansuri sunt rezultatul:

- Algoritmilor noi și a euristicilor mai eficiente.
- Tehnicilor rafinate de implementare, care s-au dovedit esențiale pentru succes.

5.1 Rularea bencmark-ului

Competiția SAT 2024 include patru secțiuni principale, fiecare oferind oportunități distincte pentru evaluarea și dezvoltarea solverelor SAT:

- **Main Track:** Solvere SAT secvențiale.

- **Cloud Track:** Solvere rulate pe infrastructură cloud.
- **No Limit Track:** Fără restricții hardware sau temporale.
- **Parallel Track:** Solvere optimizate pentru paralelism.

Pentru analiza noastră, ne-am concentrat pe Main Track, deoarece aceasta reprezintă standardul principal pentru testarea performanței solverelor SAT. În cadrul acestei secțiuni in cadrul competiei au fost folosite:

- 300 și 600 de benchmark-uri
- Fiecare benchmark are o limită de timp de 5000 de secunde.
- Resursele hardware standard includ 128 GB RAM
- Solverii din toate pistele vor fi clasați pe baza scorului PAR-2, care este suma timpului de rulare plus de două ori timpul de expirare pentru instanțe nerezolvate.

In cadrul testarii noastre:

- 30 de benchmark-uri dintr-o familie numita miter
- Fiecare benchmark are o limită de timp de 2880 de secunde(ca sa putem rula 30 in timp de 24 de ore).
- Resursele hardware standard includ 16 GB RAM.

Pentru evaluarea performanței solverului SAT MiniSat, am dezvoltat un proces structurat pentru rularea unui set de 30 de fișiere de intrare în format .cnf. Procesul este descris detaliat mai jos.

Pentru o organizare eficientă, am creat două directoare principale. Fișiere de ieșire, acestea conțin soluțiile găsite de solver pentru fiecare fișier. Fișiere de statistici, acestea includ informații despre timpul de execuție și eventualele erori.

Pentru a asigura finalizarea testelor în intervalul de 24 de ore, fiecare execuție a fost limitată la 2880 de secunde. Această limitare a fost implementată folosind comanda gtimeout. Testele au fost efectuate conform următorilor pași:

1. Scriptul a iterat prin toate fișierele .cnf din directorul de intrare.
2. Pentru fiecare fișier solver-ul MiniSat a fost rulat utilizând gtimeout pentru a respecta limita de timp, iar rezultatele au fost salvate în fișiere dedicate în fișiere de soluție (conțin rezultatul obținut ex SAT sau UNSAT, sau INDET) și în fișiere de statistici(acestea documentează durata de execuție și erorile întâlnite).

5.2 Rezultate obtinute

Rezultatele obținute în urma testelor efectuate oferă o perspectivă clară asupra performanței solver-ului MiniSat în raport cu constrângerile hardware și temporale determinate de noi. În această secțiune, vom prezenta atât statisticile cheie, cât și observațiile relevante privind timpul de execuție, numărul de soluții găsite și comportamentul în cazul fișierelor complexe. De asemenea, vom analiza impactul resurselor hardware disponibile (16 GB RAM și procesor M1 Pro)

asupra capacității de procesare a instanțelor SAT, comparativ cu timpul limită de execuție impus (2880 secunde).

În urma executării a 30 de fișiere noi am creat un tabel pentru a urmări cum se comportă Minisat în raport cu hardware care îl avem noi și mărimea și complexitatea fișierului în sine. Astfel am observat ca am observat că majoritatea fișierelor de intrare nu au reușit să fie rezolvate în limita de timp stabilită, iar solver-ul MiniSat a fost adesea nevoit să se oprească înainte de a găsi o soluție completă. Aceasta subliniază complexitatea instanțelor testate și relația directă dintre timpul de execuție și resursele hardware disponibile. În cazul fișierelor cu complexitate mare (atât în ceea ce privește numărul de variabile, cât și al clauzelor), se poate observa toate lucrurile constate, ceea ce sugerează că aceste instanțe necesită mai multe resurse pentru a fi procesate în mod eficient. Pe baza tabelu-

Table 1. Summary of MiniSat Benchmark Results

File Name	Variables	Clauses	Conflicts	Decisions	CPU Time (s)	Result
benchmarkfile1.out	62500	204018	29708704	48128071	2837.68	INDET
benchmarkfile2.out	377084	1227752	7250914	23210158	2791.62	INDET
benchmarkfile3.out	5068	15159	37386591	43015895	2860.41	INDET
benchmarkfile4.out	85224	276256	25346482	53160034	2856.89	INDET
benchmarkfile5.out	5253	15714	37665303	43499562	2853.09	INDET
benchmarkfile6.out	44195	128348	27515226	42730372	2843.85	INDET
benchmarkfile7.out	4595	13744	30014170	33877577	2046.53	UNSAT
benchmarkfile8.out	1454067	4368062	4794121	629288053	2753.86	INDET
benchmarkfile9.out	2861	8538	79919298	97135775	2854.15	INDET
benchmarkfile10.out	203492	602144	46205580	160598330	2837.88	INDET
benchmarkfile11.out	1454447	4369240	5181035	555213477	2803.24	INDET
benchmarkfile12.out	48505464	120975196	26216	239335536	1028.45	INDET
benchmarkfile13.out	35987452	35987452	54414	744156224	1505.21	INDET
benchmarkfile14.out	4538	13573	40933294	46311917	2841.62	INDET
benchmarkfile15.out	5462	16339	37199213	43346708	2844	INDET
benchmarkfile16.out	3426	10231	74515628	91602426	2817.63	INDET
benchmarkfile17.out	100238	332316	14592808	3719779	2813.94	INDET
benchmarkfile18.out	3111	9288	75407032	90345235	2795.2	INDET
benchmarkfile19.out	2780	8297	98827073	114357472	2837.41	INDET
benchmarkfile20.out	85230	276272	25323587	51335360	2816.84	INDET
benchmarkfile21.out	5189	15520	38324966	44261680	2864.38	INDET
benchmarkfile22.out	2357	7030	60321247	72009980	1786.29	UNSAT
benchmarkfile23.out	29277589	72984706	96350	353079557	1178.5	SAT
benchmarkfile24.out	4745	14194	40793160	46156597	2835.76	INDET
benchmarkfile25.out	162	774	50133511	58308211	290.156	UNSAT
benchmarkfile26.out	4842964	12045456	6163	7470746	33.7055	SAT
benchmarkfile27.out	572519	1715261	5543998	15503270	2835.99	INDET
benchmarkfile28.out	398684	1325288	8526833	33131522	2791.59	INDET
benchmarkfile29.out	49370	144360	75085653	201035009	2853.47	INDET
benchmarkfile30.out	4640	13079	32747229	46037499	614.082	UNSAT

lui, aputem constata următoarele observații referitoare la performanța solverului MiniSat în cadrul testarii bencmark-urilor:

Rezultate INDET (Indeterminate): Majoritatea fișierelor (24 din 30) au returnat rezultatul "INDET", ceea ce sugerează că MiniSat nu a reușit să determine satisfiabilitatea pentru aceste fișiere într-un timp rezonabil. Acest lucru poate fi cauzat de complexitatea mare a problemelor (numărul mare de variabile și clauze).

Rezultate UNSAT (Unsatisfiable): Au fost obținute 4 fișiere cu rezultate "UNSAT", ceea ce indică faptul că MiniSat a reușit să stabilească că aceste probleme nu au soluție.

Rezultate SAT (Satisfiable): Două fișiere au returnat "SAT", adică MiniSat a găsit o soluție satisfiabilă pentru acestea. Ca exempleu fișierul benchmarkfile26.out a fost rezolvat extrem de rapid, în doar 33.7 secunde, în timp ce fișierul benchmarkfile23.out a avut un timp de 1178.5 secunde.

Daca raportam aceste rezultate cu resursele folosite de noi(laptop pe procesor M1 Pro cu 16 GB RAM) putem spune ca SAT solvers întâmpina dificultăți în a exploata complet arhitectura multi-core a unui procesor modern, deoarece aceștia pot să nu fie complet optimizați pentru a profita de paralelismul complet pe mai multe nuclee. De asemenea, M1 Pro poate avea un impact asupra performanței în funcție de modul în care MiniSat utilizează procesele de rulare. Și astfel întrucât MiniSat nu este neapărat paralelizabil în mod eficient pe mai multe nuclee (dependența puternică de procesarea secvențială în cadrul algoritmilor SAT poate limita performanța), procesorul M1 Pro ar putea să nu fie complet exploatat. De asemenea, în cazul unui număr mare de conflicte sau decizii, performanța poate fi afectată de complexitatea intrărilor și de resursele disponibile pentru fiecare proces de calcul.

O altă observație ar fi timpul limită de execuție care noi l-am setat de a fi 2880 secunde(48 de minute) poate fi o mică problemă deoarece diferă de timpul folosit in competiții care e de 5000 secunde, astfel posibil ca timpul setat de noi nu era destul pentru Minisat pentru a rezolva unele benchmark-uri ca rezultat obținând atât de multe rezultate INDET.

În concluzie, rezultatele sugerează că MiniSat are performanțe bune în identificarea instanțelor nesatisfiabile și satisfiabile în cazurile mai simple, dar întâmpină dificultăți semnificative atunci când se confruntă cu probleme mai complexe, caracterizate printr-un număr mare de variabile și clauze. Îmbunătățirea algoritmilor și a resurselor utilizate ar putea duce la o performanță mai bună în abordarea instanțelor "INDET", în special în diferite competiții cu limitări de timp mai stricte.

References

1. GitHub Project. Available: https://github.com/Damyy17/project-vf-team-8
2. Z. Kincaid: "SAT Handbook - CDCL". [Online]. Available: https://www.cs.princeton.edu/ zkincaid/courses/fall18/readings/SATHandbook-CDCL.pdf, last accessed 2024/11/10
3. J. Worrell: "Lecture Notes - Lecture 6". [Online]. Available: https://www.cs.ox.ac.uk/james.worrell/lecture06.pdf, last accessed 2024/11/10
4. Minisat Page. [Online]. Available: https://http://minisat.se/, last accessed 2024/11/13
5. Sat Competition 2024. [Online]. Available: https://satcompetition.github.io/2024/tracks.html, last accessed 2024/11/12
6. A. Biere: "VTSA 2012 Talk: SAT and Model Checking". [Online]. Available: https://resources.mpi-inf.mpg.de/departments/rg1/conferences/vtsa12/slides/biere/Biere-VTSA12-talk.pdf, last accessed 2024/11/14
7. "MiniSat Documentation." [Online]. Available: https://www.decision-procedures.org/handouts/MiniSat.pdf, last accessed 2024/11/14
8. "Model Checking Lecture Notes." [Online]. Available: https://swtv.kaist.ac.kr/courses/cs453-fall09/model_checkin_2.pdf, last accessed 2024/11/14

Authors' Contributions

Grosu Damian: Secțiunea 5, "Benchmark SAT Competition", Raportul in Latex.
Scoropad Iulian: Secțiunea 4, "Algoritmii DPLL și CDCL"
Daniel Dobrescu: Secțiunea 3, "Minisat"
Anton Constantin-Adrian: Secțiunea 1 si 2, "Introducere" și "Descrierea problemei"