

KOP úloha 1

Jorge Zuñiga

zunigjor

2022

Executive summary	1
Pilot	1
Experiment	1
Metriky a výsledky	1
Závěr	1
Úvod	2
Algoritmy GSAT a ProbSAT	2
Sady instancí	2
Pilotní experiment	3
Hlavní experiment	4
Metriky	4
Výsledky	4
Grafy výsledků	5
succ	5
steps	6
avg fined	7
mu	8
sig ²	9
winner	10
Závěr	11
Zdroje	11

Executive summary

Práce se věnovala porovnání algoritmů GSAT a ProbSAT při řešení problémů 3-SAT.

Pilot

Pilotní experiment byl proveden na menším vzorku instancí z datových sad uf20-91, uf50-200 a uf75-325, s menším počtem 100 běhů a sloužil hlavně k určení proměnné max-flips.

Výsledkem pilotu byl výběr proměnné max-flips na hodnotu 1500.

Experiment

Ostrý experiment proběhl nad celými sadami instancí z datových sad uf20-91, uf50-200 a uf75-325 a statisticky významnějším vyšším počtem 1000 běhů.

Metriky a výsledky

V rámci experimentu byly zpracovány všechny metriky z [3], pro finální rozhodnutí byly vybrány metriky:

- succ** Metrika počtu úspěšných běhů u sad uf20-91 a uf50-200 nenabízí moc rozdílů k rozhodnutí mezi algoritmy.
U sady uf75-325 je ale situace znatelně jiná, z histogramu je jasné vidět **dominance algoritmu ProbSAT**. U této metriky je vyšší počet úspěšných kroků lepší.
- steps** Metrika průměrného počtu flipů k vyřešení u všech sad opět **naznačuje dominanci ProbSATu**. U této metriky je nižší počet kroků lepší.
- avg fined** Metrika penalizovaného průměrného počtu běhů napříč sadami opět indikuje ve **prospěch ProbSATu**. U této metriky je nižší počet kroku lepší.
- mu a sig²** Metriky odhadů parametrů lognormálního rozdělení napříč datovými sadami **jemně naznačují ve prospěch algoritmu ProbSAT**. Tedy, že algoritmus ProbSAT s vyšší pravděpodobností skončí v menším počtu kroků než GSAT.
- winner** Metrika určující algoritmus, jehož korigovaná CDF leží po posledním průsečíku výše napříč všemi sadami **jasně naznačuje dominanci ProbSATu**.

Závěr

Na základě výsledků z vybraných metrik, je ProbSAT lepším a efektivnějším algoritmem než GSAT, pro řešení problémů 3-SAT.

Úvod

Práce se zabývá porovnáním algoritmů GSAT a ProbSAT při řešení problémů 3-SAT.

Algoritmy GSAT a ProbSAT

Cíl obou algoritmů je najít ohodnocení proměnných, tak aby byla formule v konjunktivní normální formě byla splněna.

Algoritmus GSAT byl převzat z courses stránek předmětu NI-KOP [1] a byl nastaven parametr $p=0.4$. Algoritmus ProbSAT byl převzat z [2] a upraven pro potřeby experimentu. Parametry byly nastaveny na $cb=2.3$ a $cm=0$. U obou byl nastaven přepínač seedu randomizace `-r` na hodnotu `time`.

Výstupem obou algoritmů pro instanci problému je čtveřice hodnot: počet kroků algoritmu, max. počet kroků algoritmu, počet splněných klauzulí, max. počet splněných klauzulí.

Sady instancí

Algoritmy byly zkoumány nad datovými sadami `uf20-91`, `uf50-200` a `uf75-325` z [1]. První číslo v názvu označuje počet proměnných v instancích. Sady obsahují postupně 1000, 1000 a 100 instancí.

Pilotní experiment

Cílem pilotního experimentu bylo vybrat vhodnou hodnotu maximálního počtu flipů.

Pro pilotní experiment bylo vybráno prvních 100 instancí z každé sady a na každé bylo spuštěno 100 běhů.

Testované hodnoty flipů byly: 300, 700, 1500, 3000, 6000. Následující tabulky obsahují výsledky měřených metrik.

GSAT

Flipy	Celkem běhů	Splněno	Dosaženo max flipů	Splněno / flipů
300	30 000	20 500	9 510	68.33
700	30 000	22 993	7 013	32.85
1500	30 000	24 905	5 096	16.60
3000	30 000	26 401	3 600	8.80
6000	30 000	27 685	2 316	4.61

Průměrně splněno	24 496.80
Průměrně splněno / flipů	26.24

ProbSAT

Flipy	Celkem běhů	Splněno	Dosaženo max flipů	Splněno / flipů
300	30 000	21 826	8181	72.75
700	30 000	24 726	5279	35.32
1500	30 000	26 797	3205	17.86
3000	30 000	28 321	1680	9.44
6000	30 000	29 337	663	4.89

Průměrně splněno	26 201.40
Průměrně splněno / flipů	28.05

Z výsledků je vidět, že čím více flipů, tím více se zvyšuje počet úspěšně splněných instancí. Ovšem je důležité uvědomit si, že více flipů znamená delší běh a zvyšování flipů nezvyšuje počet splněných instancí lineárně.

Proto byl spočítán průměr z těchto hodnot a průměr hodnot splněno / flipů a bylo **pro hodnotu maximálního počtu flipů bylo vybráno 1500 flipů.**

Hlavní experiment

Hlavní experiment proběhl nad všemi instancemi v sadách uf20-91, uf50-200 a uf75-325 a pro každou instanci bylo spuštěno 1000 běhů. Celkem bylo spuštěno 2 100 000 běhů pro každý z algoritmů.

Maximální počet flipů byl nastaven na 1500.

Metriky

Pro všech 2 100 000 výsledků pro každý z algoritmů (1000 běhů pro každou instanci) byly zpracovány metriky podle [3].

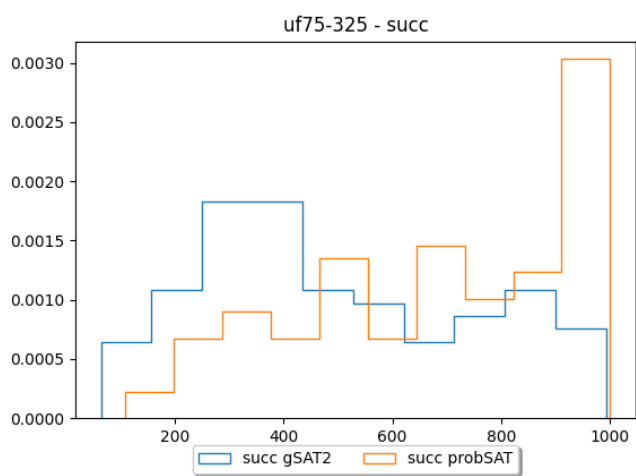
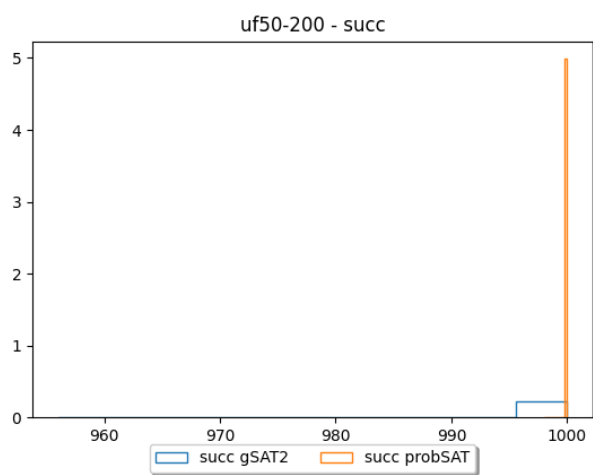
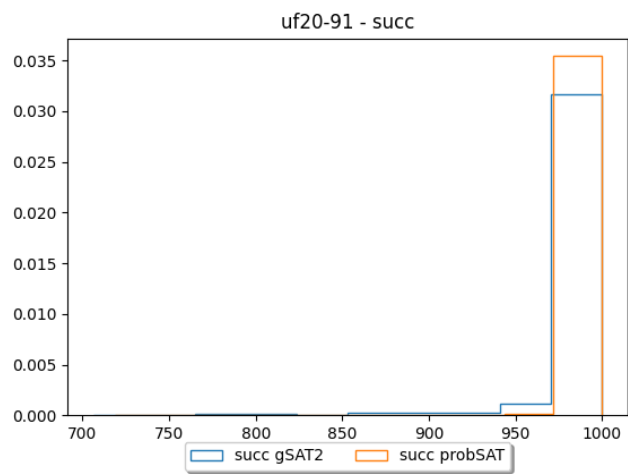
K vyhodnocení experimentu jsem se zaměřil hlavně metriky succ, steps, awg fined, mu, sig² a winner.

Výsledky

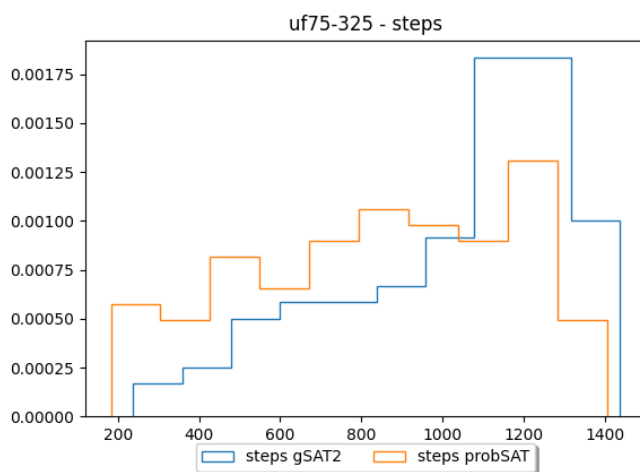
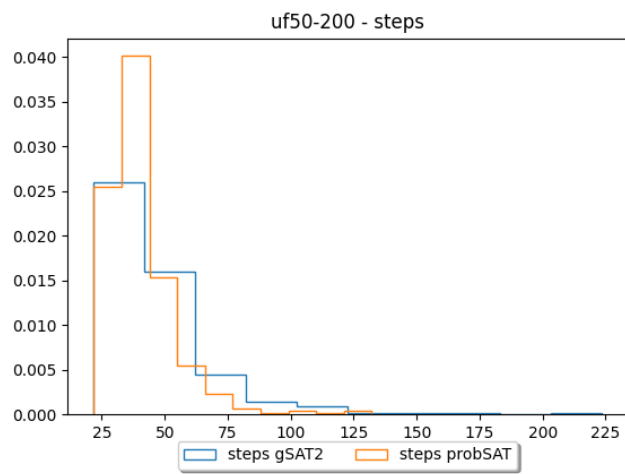
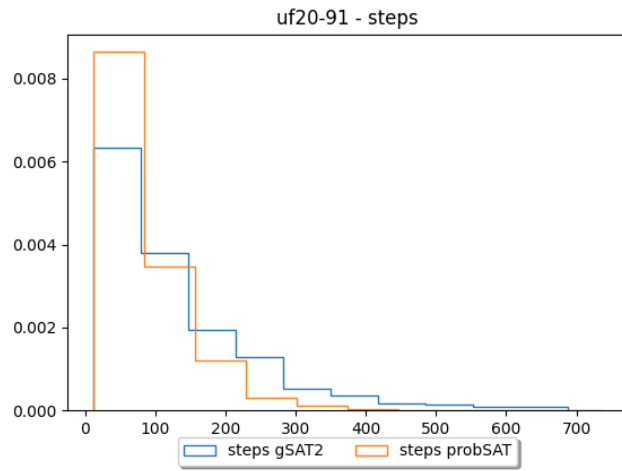
- succ** Metrika počtu úspěšných běhů u sad uf20-91 a uf50-200 nenabízí moc rozdílů k rozhodnutí mezi algoritmy.
U sady uf75-325 je ale situace znatelně jiná, z histogramu je jasné vidět dominance algoritmu ProbSAT. U této metriky je vyšší počet úspěšných kroků lepší.
- steps** Metrika průměrného počtu flipů k vyřešení u všech sad opět naznačuje dominanci ProbSATu. U této metriky je nižší počet kroků lepší.
- awg fined** Metrika penalizovaného průměrného počtu běhů napříč sadami opět indikuje ve prospěch ProbSATu. U této metriky je nižší počet kroku lepší.
- mu a sig²** Metriky odhadů parametrů lognormálního rozdělení napříč datovými sadami jemně naznačují ve prospěch algoritmu ProbSAT. Tedy, že algoritmus ProbSAT s vyšší pravděpodobností skončí v menším počtu kroků než GSAT.
- winner** Metrika určující algoritmus, jehož korigovaná CDF leží po posledním průsečíku výše napříč všemi sadami jasné naznačuje dominanci ProbSATu.

Grafy výsledků

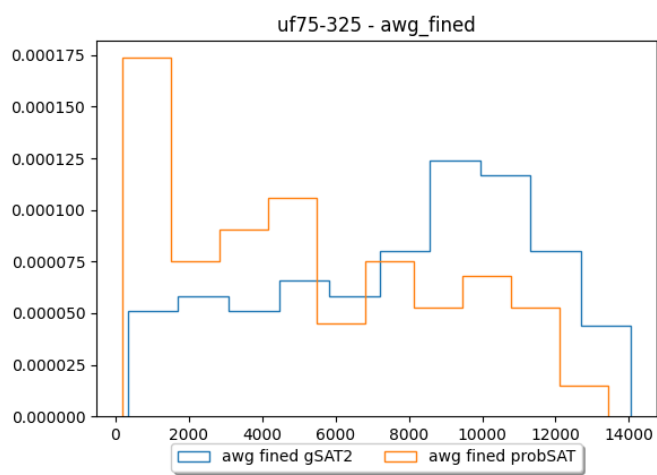
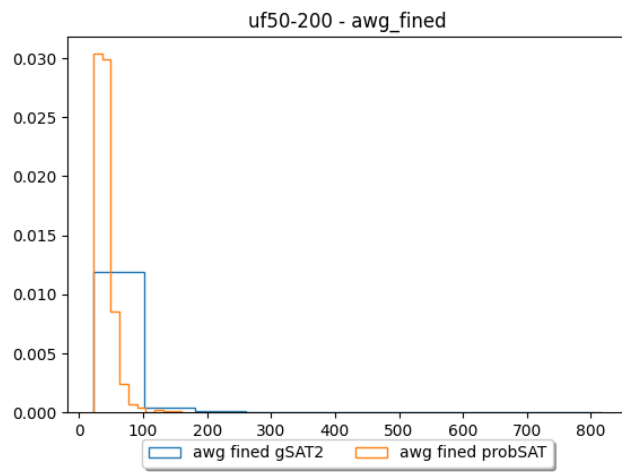
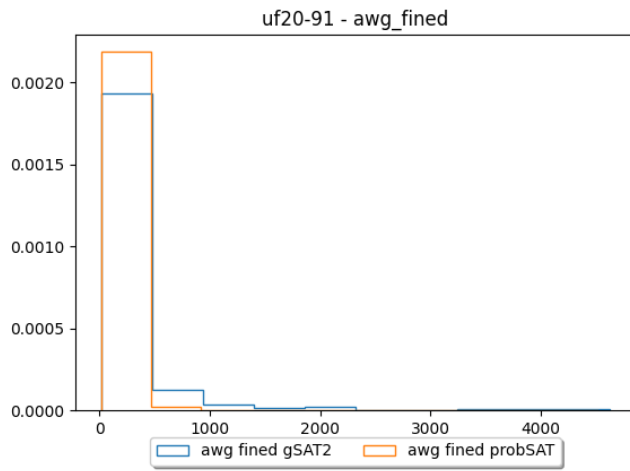
succ



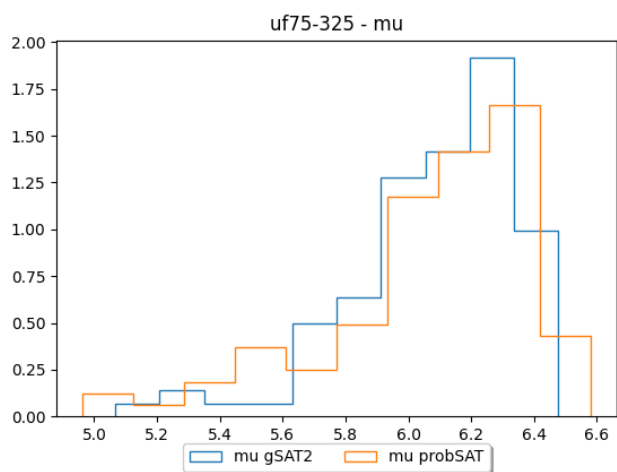
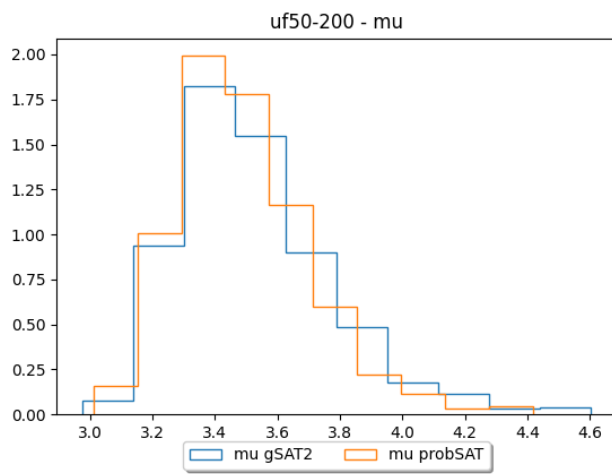
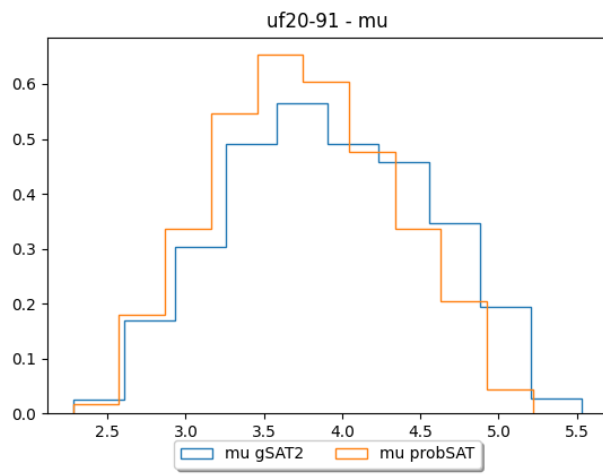
steps



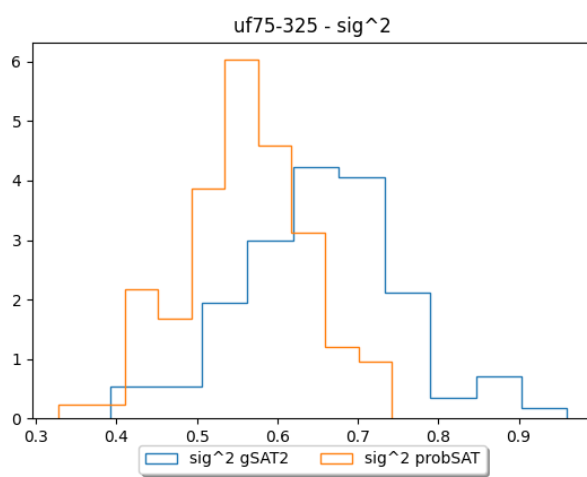
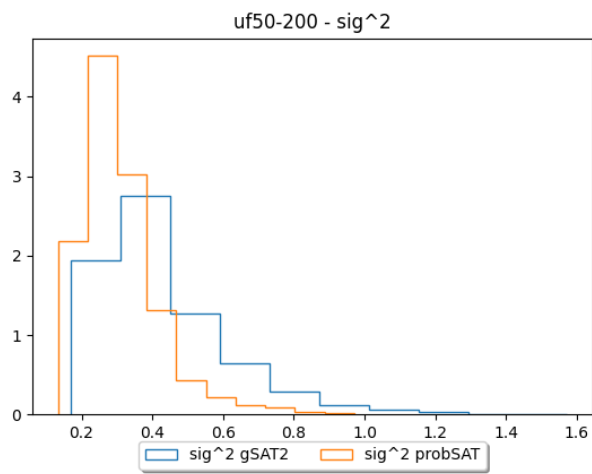
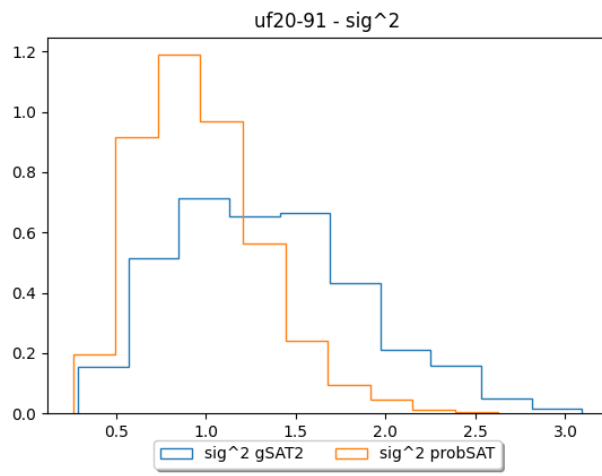
avg fined



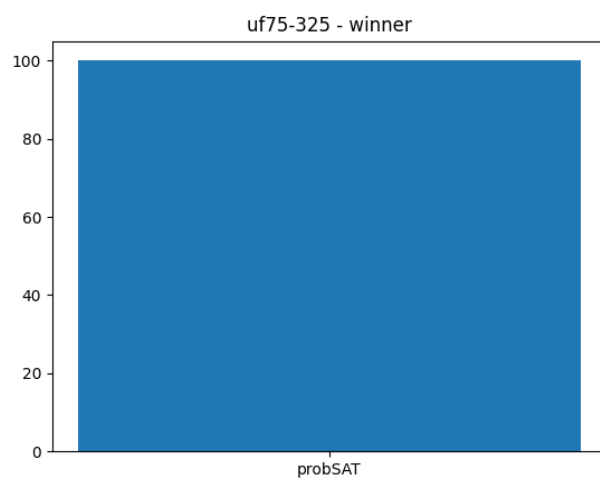
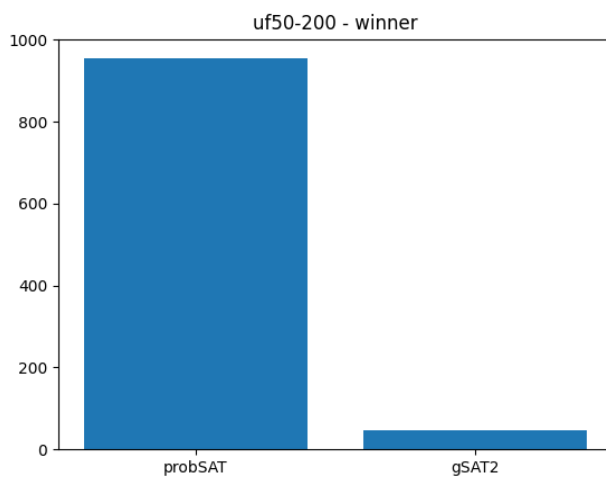
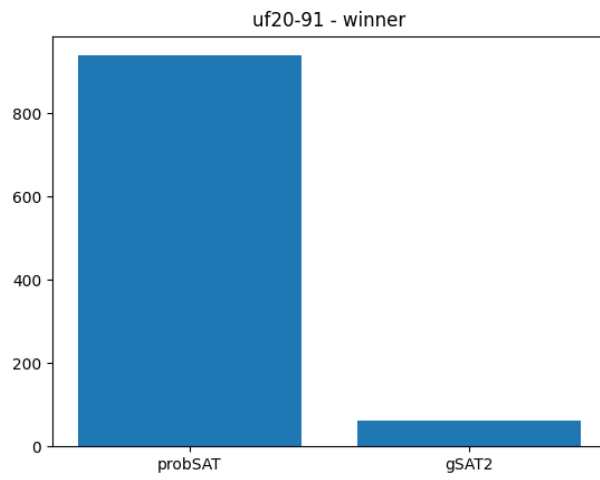
mu



sig²



winner



Závěr

Na základě výsledků z vybraných metrik, je ProbSAT lepším a efektivnějším algoritmem než GSAT, pro řešení problému 3-SAT.

Zdroje

- [1] Jan Schmidt. Data a programy. url: <https://courses.fit.cvut.cz/NI-KOP/download/index.html>
- [2] Adrian Balint. The probsat sat solver. url: <https://github.com/adrianopolus/probSAT>
- [3] Jan Schmidt. Data pro srovnání algoritmů. url: <https://courses.fit.cvut.cz/NI-KOP/tutorials/files/data-comp.html>