



4 KWIETNIA 2020

**ZADANIE 2: PROBLEMY SPEŁNIANIA
OGRANICZEŃ**
SPRAWOZDANIE

AGNIESZKA MIECHOWICZ, 242533



Sudoku 42 – mieszane 1 (40 rozwiązań):

8	7	-	-	-	-	-	-	9
-	-	-	-	-	5	8	-	-
-	-	3	6	4	-	-	-	-
-	-	-	-	-	4	7	-	5
-	-	-	-	-	-	2	-	-
-	2	-	-	-	1	-	-	-
-	-	6	-	2	-	9	-	-
-	8	4	-	-	-	-	-	-
-	9	-	5	-	-	-	-	-

Sudoku 45 – mieszane 2 (0 rozwiązań):

8	5	-	1	-	5	-	-	-
-	-	-	-	-	-	-	2	-
-	-	4	9	-	-	3	-	-
-	-	-	-	7	-	-	3	-
-	-	1	-	4	8	-	-	-
-	8	-	-	-	-	9	-	2
7	-	6	-	-	-	-	-	-
-	4	-	7	-	-	2	9	-
2	9	8	-	-	-	-	6	3

Część pierwsza: Badanie metod rozwiązywania problemów CSP dla Sudoku

Wprowadzenie

Cel badania: Porównanie ze sobą dwóch zaimplementowanych metod rozwiązywania CSP – Backtrackingu oraz Forward Checkingu pod względem czasu, liczby odwiedzonych węzłów oraz liczby nawrotów.

Stale:

Pierwsza część badania:

Heurystyka wyboru zmiennej – w kolejności definicji

Heurystyka wyboru wartości – w kolejności definicji

Druga część badania:

Heurystyka wyboru zmiennej – w kolejności definicji

Heurystyka wyboru wartości – w kolejności losowej

Zmienne w badaniu:

Metoda rozwiązywania problemu

Przebieg badania:

Dla każdego z badanych przypadków – problemów sudoku wymienionych powyżej – dziesięciokrotnie uruchomiłam algorytm Backtrackingu, a następnie dziesięciokrotnie Forward Checkingu dla jednej z par z heurystyk, a następnie dla drugiej.

Dla każdej z tych metod zapisałam średnią liczbę węzłów, nawrotów i czas, a także ich wartości przy znalezieniu pierwszego rozwiązania.

Ważne uwagi:

Przy wykorzystaniu pierwszej z podanych grup heurystyk algorytm staje się deterministyczny – przy każdym uruchomieniu daje takie same wyniki liczby węzłów oraz liczby nawrotów. Różny jest jedynie w niektórych przypadkach czas uruchomienia. (Głównie ze względu na pracę komputera, a nie sam algorytm).

Co może zaburzyć wyniki badania?

- a) W przypadku heurystyki wartości w kolejności definicji, wartości są sortowane (ponieważ dla Forward Checkingu inaczej nie byłyby w odpowiedniej kolejności). Dla Backtrackingu w zasadzie to sortowanie nie byłoby potrzebne, a więc jest to dodatkowy czynnik świadczący o jego przewadze.
- b) W przypadku losowego wyboru wartości – dla Backtrackingu czas jest dodatkowo wydłużony, ponieważ kosztowne losowanie kolejności przeprowadza się dynamicznie w każdym wywołaniu. Przy Backtrackingu zmienia się kolejność całej listy możliwych wartości, a w Forward Checkingu aktualnej dziedziny.

Porównanie metod rozwiązywania problemów dla Sudoku nr 15 – problemu łatwego

Heurystyka wyboru wartości – w kolejności definicji:

Algorytm	Węzły 1. rozw	Nawroty 1. rozw	Czas 1. rozw[ms]	Węzły	Nawroty	Czas[ms]	L.rozw
BT	2008	466	0,2	9068	2339	1,4	1
FC	306	149	0,1	1489	986	2,5	1

Tabela 1 Uśrednione parametry wykonania po dziesięciokrotnym uruchomieniu każdej z metod dla heurystyki wyboru zmiennej w kolejności definicji i heurystyki wartości w kolejności definicji dla Sudoku nr 15.

Heurystyka wyboru wartości –w kolejności losowej:

Algorytm	Węzły 1. rozw	Nawroty 1. rozw	Czas 1. rozw[ms]	Węzły	Nawroty	Czas[ms]	L.rozw
BT	4953,1	1224,9	4,1	9067,7	2339	7,3	1
FC	757,2	448,2	1,6	1489	986	4,1	1

Tabela 2 Tabela 3 Uśrednione parametry wykonania po dziesięciokrotnym uruchomieniu każdej z metod dla heurystyki wyboru zmiennej w kolejności definicji i heurystyki wyboru wartości w kolejności losowej dla Sudoku nr 15.

Dla heurystyki z definicji można zauważyć, że pomimo znacznego zmniejszenia liczby węzłów oraz nawrotów, zarówno czas wykonania algorytmu jak i czas do otrzymania pierwszego rozwiązania jest niemal *dwukrotnie dłuższy*. W tym przypadku korzyść z mniejszej złożoności obliczeniowej jednych odwiedzin węzła przewyższa tę związaną z mniejszą liczbą odwiedzonych węzłów.

Dla mniej optymalnej heurystyki – heurystyki losowej można jednak zauważyć ogromne różnice w czasie wykonania algorytmu na korzyść Forward Checkingu. Nie da się jednak ukryć, że może to być w dużej mierze spowodowane złożoną procedurą losowania, która dla Backtrackingu jest wielokrotnie powtarzana na całej dziedzinie, a nie tylko jej części.

Można zauważyć, że metody mają różny stosunek nawrotów do odwiedzonych węzłów. Nie jest zaskoczeniem, że w przypadku Forward Checkingu liczba nawrotów stanowi ponad połowę liczby węzłów, a dla Backtrackingu tylko około 0,26.

Porównanie metod rozwiązywania problemów dla Sudoku nr 26– problemu średniego

Heurystyka wyboru wartości – w kolejności definicji:

Algorytm	Węzły do 1	Nawroty do 1	Czas do 1 [ms]	Węzły	Nawroty	Czas[ms]	Liczba rozw
BT	733156	208640	106,8	826944	236104	120,2	1
FC	91977	60347	115,5	93519	61407	117,5	1

Tabela 4 Uśrednione parametry wykonania po dziesięciokrotnym uruchomieniu każdej z metod dla heurystyki wyboru zmiennej w kolejności definicji i heurystyki wartości w kolejności definicji dla Sudoku nr 26.

Heurystyka wyboru wartości – w kolejności losowej:

Algorytm	Węzły do 1	Nawroty do 1	Czas do 1 [ms]	Węzły	Nawroty	Czas[ms]	Liczba rozw
BT	436053,8	124368,1	341,7	826944,8	236104	648,8	1
FC	40826,8	26781	110,4	93519	61407	253,1	1

Tabela 5 Uśrednione parametry wykonania po dziesięciokrotnym uruchomieniu każdej z metod dla heurystyki wyboru zmiennej w kolejności definicji i heurystyki wyboru wartości w kolejności losowej dla Sudoku nr 26.

Pomimo trudności zaledwie średniej, jest to najbardziej złożone obliczeniowo sudoku spośród wybranych. Można zauważyć, że dla tego problemu czas wykonywania Backtrackingu jest niewiele dłuższy niż czas Forward Checkingu. Jest to różnica zaledwie kilku milisekund, jest jednak wyraźna.

Tym razem liczba odwiedzonych węzłów jest niemal dziewięciokrotnie większa dla Backtrackingu niż dla Forward Checkingu.

Porównanie metod rozwiązywania problemów dla Sudoku nr 36 – problemu trudnego

Heurystyka wyboru wartości – w kolejności definicji:

Algorytm	Węzły do 1	Nawroty do 1	Czas do 1 [ms]	Węzły	Nawroty	Czas[ms]	L.rozw
BT	54253	14566	8,5	65473	17608	10,5	1
FC	6533	4349	8,2	7882	5277	10,1	1

Tabela 6 Uśrednione parametry wykonania po dziesięciokrotnym uruchomieniu każdej z metod dla heurystyki wyboru zmiennej w kolejności definicji i heurystyki wyboru wartości w kolejności definicji dla Sudoku nr 36.

Heurystyka wyboru wartości – w kolejności losowej:

Algorytm	Węzły 1. rozw	Nawroty 1. rozw	Czas 1. rozw[ms]	Węzły	Nawroty	Czas[ms]	L.rozw
BT	25118,8	6710,9	18,5	65469,7	17608	48,6	1
FC	4555,6	3012,9	12,3	7882	5277	21,7	1

Tabela 7 Uśrednione parametry wykonania po dziesięciokrotnym uruchomieniu każdej z metod dla heurystyki wyboru zmiennej w kolejności definicji i heurystyki wyboru wartości w kolejności losowej dla Sudoku nr 36.

Dla tego problemu widać nieznaczną poprawę czasu (dosłownie o ułamki milisekund) Forward Checkingu w stosunku do Backtrackingu.

Można też zauważyć podobną do poprzedniego przypadku zależność - dziewięciokrotnie mniejszą liczbę odwiedzonych węzłów. 3,08

Porównanie metod rozwiązywania problemów dla Sudoku nr 42– problemu mieszanego 1

Heurystyka wyboru wartości – w kolejności definicji:

Algorytm	Węzły 1. rozw	Nawroty 1. rozw	Czas 1. rozw[ms]	Węzły	Nawroty	Czas[ms]	L.rozw
BT	26042	5482	3	459655	98015	63,1	40
FC	2027	1411	2,2	46808	33789	57,5	40

Tabela 8 Uśrednione parametry wykonania po dziesięciokrotnym uruchomieniu każdej z metod dla heurystyki wyboru zmiennej w kolejności definicji i heurystyki wartości w kolejności definicji dla Sudoku nr 42.

Heurystyka wyboru wartości – w kolejności losowej:

Algorytm	Węzły 1. rozw	Nawroty 1. rozw	Czas 1. rozw[ms]	Węzły	Nawroty	Czas[ms]	L.rozw
BT	47155,5	10064,9	29,5	459665,3	98015	289,8	40
FC	4710,1	3326,1	12,6	46808	33789	131	40

Tabela 9 Uśrednione parametry wykonania po dziesięciokrotnym uruchomieniu każdej z metod dla heurystyki wyboru zmiennej w kolejności definicji i heurystyki wyboru wartości w kolejności losowej dla Sudoku nr 36

W przypadku tego problemu nieco wyraźniej widać różnicę czasu działania algorytmu na korzyść Forward Checkingu. Liczba odwiedzonych węzłów jest niemal dziesięciokrotnie mniejsza.

Porównanie metod rozwiązywania problemów dla Sudoku nr 45– problemu mieszanego2

Heurystyka wyboru wartości – w kolejności definicji:

Algorytm	Węzły 1. rozw	Nawroty 1. rozw	Czas 1. rozw[ms]	Węzły	Nawroty	Czas[ms]	L.rozw
BT	-	-	-	46265	10913	6,2	0
FC	-	-	-	8322	6531	11,1	0

Tabela 10 Uśrednione parametry wykonania po dziesięciokrotnym uruchomieniu każdej z metod dla heurystyki wyboru zmiennej w kolejności definicji i heurystyki wartości w kolejności definicji dla Sudoku nr 45.

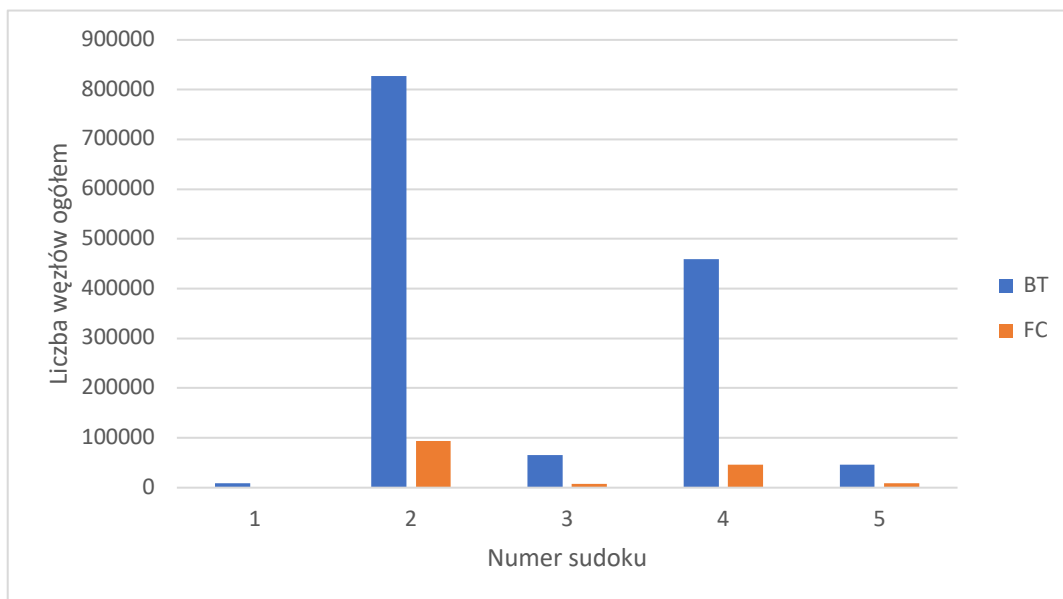
Heurystyka wyboru wartości – w kolejności losowej:

Algorytm	Węzły 1. rozw	Nawroty 1. rozw	Czas 1. rozw[ms]	Węzły	Nawroty	Czas[ms]	L.rozw
BT	-	-	-	46265	10913	31,1	0
FC	-	-	-	8322	6531	26,1	0

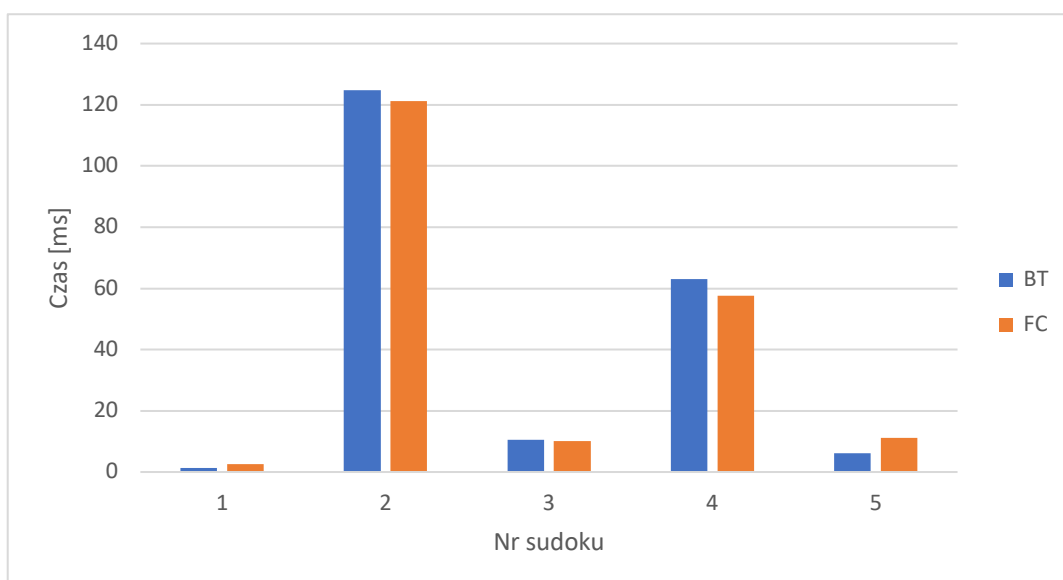
Tabela 11 Uśrednione parametry wykonania po dziesięciokrotnym uruchomieniu każdej z metod dla heurystyki wyboru zmiennej w kolejności definicji i heurystyki wyboru wartości w kolejności losowej dla Sudoku nr 36

W przypadku tego problemu da się zauważyć ogromną różnicę czasów wykonania algorytmu na korzyść backtrackingu.

Różnica w liczbie odwiedzonych węzłów jest jedynie pięciokrotna na korzyść Forward Checkingu.



Wykres 1 Liczba węzłów dla każdej z metod dla poszczególnych problemów.



Wykres 2 Czas wykonania każdej z metod dla poszczególnych problemów.

Wnioski:

Dla **Forward Checkingu** obserwujemy **znaczące, wielokrotne zmniejszenie liczby odwiedzonych węzłów oraz nawrotów**. W wielu przypadkach dla liczby węzłów jest to różnica około dziewięciokrotna na korzyść Forward Checkingu w stosunku do Backtrackingu. Może być to związane z wielkością dziedzin oraz liczbą ograniczeń w sudoku.

Nie idzie to jednak w parze ze znaczącym skróceniem działania metody. Dla problemu łatwego jest on nawet wydłużony. Dla problemów – średniego, trudnego oraz mieszanego jest nieco skrócony. Można zauważyć, że w przypadku problemu trwającego najdłużej, różnica w czasie działania algorytmów zwiększa się, czyli jest proporcjonalna do czasu ogółem.

Im większy stosunek liczby węzłów Backtrackingu do Forward Checkingu, tym większa różnica czasu na korzyść tego drugiego. **Można zauważyć, że przy około dziewięciokrotnej różnicy, czas Forward Checkingu zaczyna być lepszy.** Być może jest to związane ze złożonością obliczeniową sprawdzania ograniczeń w przód (które to jest związane z wielkością dziedziny/ograniczeń). Można wyciągnąć wniosek, że złożoność obliczeniowa dla pojedynczego węzła w Backtrackingu jest około 9-krotnie mniejsza niż w Forward Checkingu. Gdy węzłów jest więcej niż dziewięciokrotnie mniej, Forward Checking staje się rozwiązaniem szybszym. Hipoteza ta wymaga jednak dalszych badań.

Dodatkowo w dwóch przypadkach czas działania Backtrackingu jest znacznie lepszy niż Forward Checkingu – dla prostego algorytmu oraz dla algorytmu bez rozwiązań.

Stosunek liczby nawrotów do liczby węzłów różni się dla obu metod. W przypadku Forward Checkingu wynosi na ogół więcej niż 0.5, natomiast dla Backtrackingu jest to tylko około 0.25. Ta zależność nie jest zaskakująca – zmniejszona liczba węzłów przy przeszukiwaniu w przód jest związana z mniejszą liczbą wartości w dziedzinie. Nawroty w naturalny sposób występują więc częściej.

Stosunek liczby węzłów odwiedzonych do 1. rozwiązania do liczby węzłów ogółem jest dla obu metod niemal identyczny (z dokładnością do 2 miejsc po przecinku).

Badanie dwóch różnych heurystyk wyboru wartości pokazało mocne i słabe strony obu algorytmów, ale także wpływ implementacji na badania. Przede wszystkim dało się zauważyć ogromną przewagę Forward Checkingu, jeśli chodzi o czas wykonania algorytmu dla losowego wyboru wartości, ale prawdopodobną przyczyną nie jest wcale zwiększona liczba węzłów, lecz implementacja samej metody – Backtracking mając w każdym węźle tę samą, dużą dziedzinę w naturalny sposób jest na przegranej pozycji. Będzie to prawdziwe również dla innych heurystyk, dla których złożoność obliczeniowa silnie zależy od dziedziny – np. dla heurystyki wyboru zmiennej w zależności od rozmiaru dziedziny.

Dodatkowe badanie:

Ponieważ zauważyłam, że stosunek czasów FC i BT jest silnie zależny od stosunku liczby odwiedzonych węzłów, postanowiłam dodatkowo zbadać tę zależność.

Cel badania: Porównanie czasów wykonania, liczby węzłów oraz zależności między nimi dla Forward Checkingu oraz Backtrackingu dla różnych problemów w zależności od ich stopnia trudności.

Stałe:

Heurystyka wyboru zmiennej – w kolejności definicji

Heurystyka wyboru wartości – w kolejności definicji

Zmienne w badaniu:

Metoda rozwiązywania problemu

Przebieg badania:

Dla wszystkich dostępnych sudoku dziesięciokrotnie uruchomiłam algorytm Backtracking oraz Forward Checking. Średnią parametrów określonych w zadaniu zapisałam do tabeli.

Wyniki badania pogrupowane według kategorii:

Trudność	Średnie węzły BT	Średnie węzły FC	CzasFC/CzasBT	WezłyBT/WezłyFC	Ile lepszych FC	Liczba sudoku
0	696053,00	35509,20	0,69	14,91	4	5
1	512998,20	41556,60	0,72	13,20	5	5
2	82072,40	10951,60	1,29	8,43	2	5
3	353174,00	53431,60	1,14	8,35	2	5
4	1092300,64	70297,00	0,63	14,32	5	5
5	337052,80	37886,80	0,97	9,23	3	5
6	1080801,76	62774,20	0,72	16,21	3	5
7	976167,40	123402,00	1,08	7,99	0	5
8	548567,33	47975,33	0,75	11,60	3	3
9	308872,00	15724,33	1,33	12,38	1	3
Suma końcowa	613596,07	51524,87	0,92	11,63	28	46

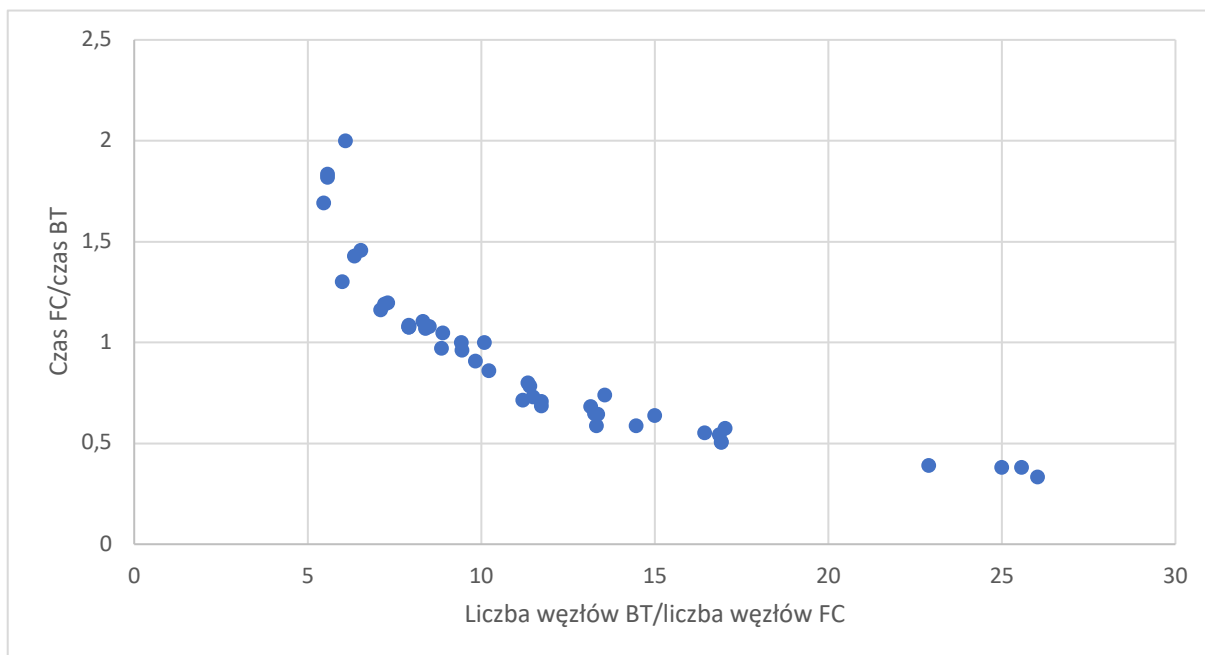
Tabela 12 Średnia liczba węzłów dla poszczególnych metod, średni stosunek czasów FC/BT oraz węzłów BT/FC oraz liczba sudoku, które są lepiej rozwiązywane przez Forward Checking niż Backtracking w zależności od kategorii problemów.

Tabela pokazuje, że w mojej implementacji wyniki nie są w silny sposób zależne od „trudności” –jedynie w niewielkim stopniu widać różnicę, jeśli chodzi o liczbę węzłów. Można jednak zauważyć, że dla niektórych grup wyraźnie odznacza się liczba sudoku, dla których lepszy jest Forward Checking. Okazuje się, że dla kategorii mieszane 1 (z wieloma rozwiązaniami) algorytm Forward Checking sprawdza się bardzo dobrze, a dla np. trudności 7 (trudnej) dla żadnego z sudoku nie dał lepszych wyników niż Forward Checking.

Warto jednak mieć świadomość, że liczba sudoku 3-5 to zbyt mało na wyciąganie dalekosiężnych wniosków, choć może to być pewną wskazówką.

IdProbl	Trudnosc	CzasBT	WezlyBT	Czas FC	WezlyFC	CzasFC/BT	WezlyBT/FC	Lepszy FC
1	0	34,6	229838	25,6	16950	0,74	13,56	1
2	0	143,6	996564	79,0	60633	0,55	16,44	1
3	0	14,8	91274	15,8	10879	1,07	8,39	0
4	0	7,0	52081	5,0	4655	0,71	11,19	1
5	0	314,0	2110508	119,2	84429	0,38	25,00	1
6	1	2,0	14726	2,0	1562	1,00	9,43	1
7	1	91,2	627463	58,2	41878	0,64	14,98	1
8	1	23,4	153907	11,8	9097	0,50	16,92	1
9	1	93,0	623249	54,4	43086	0,58	14,47	1
10	1	165,6	1145646	142,2	112160	0,86	10,21	1
11	2	2,0	16847	2,0	1670	1,00	10,09	1
12	2	21,2	140002	12,4	10522	0,58	13,31	1
13	2	11,6	83230	13,8	11536	1,19	7,21	0
14	2	23,2	161215	39,2	29541	1,69	5,46	0
15	2	1,0	9068	2,0	1489	2,00	6,09	0
16	3	69,2	447061	90,0	74539	1,30	6,00	0
17	3	13,0	89254	10,2	7831	0,78	11,40	1
18	3	17,8	127803	13,0	11133	0,73	11,48	1
19	3	149,2	1050255	212,8	165762	1,43	6,34	0
20	3	7,0	51497	10,2	7893	1,46	6,52	0
21	4	34,0	245078	18,4	14540	0,54	16,86	1
22	4	25,2	188430	17,2	14340	0,68	13,14	1
23	4	14,0	96879	9,0	7253	0,64	13,36	1
24	4	614,8	4117159	311,4	243565	0,51	16,90	1
25	4	115,8	813957	92,4	71787	0,80	11,34	1
26	5	123,4	826944	119,6	93519	0,97	8,84	1
27	5	36,8	258578	35,4	27420	0,96	9,43	1
28	5	10,2	69912	7,0	5964	0,69	11,72	1
29	5	18,4	116533	22,0	15976	1,20	7,29	0
30	5	61,0	413297	63,8	46555	1,05	8,88	0
31	6	34,8	229742	40,4	32366	1,16	7,10	0
32	6	98,8	707367	56,6	41579	0,57	17,01	1
33	6	485,2	3529821	189,2	154173	0,39	22,90	1
34	6	87,8	624613	94,8	73527	1,08	8,50	0
35	6	43,2	312466	16,4	12226	0,38	25,56	1
36	7	9,6	65473	10,6	7882	1,10	8,31	0
37	7	170,0	1203841	183,4	152282	1,08	7,91	0
38	7	171,8	1203841	184,8	152282	1,08	7,91	0
39	7	170,4	1203841	184,6	152282	1,08	7,91	0
40	7	171,0	1203841	183,8	152282	1,07	7,91	0
41	8	60,0	411890	38,8	31076	0,65	13,25	1
42	8	61,8	459655	56,0	46808	0,91	9,82	1
43	8	121,0	774157	85,4	66042	0,71	11,72	1
44	9	128,8	844789	43,0	32460	0,33	26,03	1
45	9	6,0	46265	11,0	8322	1,83	5,56	0
46	9	4,4	35562	8,0	6391	1,82	5,56	0

Tabela 13 Uśrednione parametry dziesięciokrotnego wykonania algorytmów, stosunki węzłów i czasów wykonania FC i BT dla poszczególnych sudoku na całym zbiorze dostępnych problemów.



Wykres 3 Zależność stosunku czasów wykonania poszczególnych metod od stosunku liczby węzłów dla danej metody dla heurystyki wyboru zmiennej i wartości w kolejności definicji.

Wnioski:

Badanie potwierdziło poprzednie obserwacje – im większy stosunek liczby odwiedzonych węzłów, tym lepszy czas na korzyść Forward Checkingu. Można też zauważyć, że w okolicach 9 stosunek czasów wynosi około 1. Stąd zasadny wydaje się wniosek, że średni czas wykonania dla 1 węzła w Forward Checkingu jest dziewięciokrotnie dłuższy niż w Backtrackingu.

Dla 28/46 przypadków czas wykonania metody Forward Checking był lepszy, a średni stosunek czasów FC/BT to 0.92.

Część druga: Badanie wpływu heurystyk na wynik działania algorytmu

Wprowadzenie

Cel badania:

Zbadanie wpływu heurystyk wyboru zmiennej oraz heurystyk wyboru wartości na działanie algorytmu.

Stale:

Część 1: Metoda wyboru zmiennej dla badania heurystyk wyboru wartości – w kolejności definicji

Część 2: Metoda wyboru wartości dla badania heurystyk wyboru zmiennej – w kolejności definicji

Zmienne w badaniu:

Część 1: Heurystyka wyboru zmiennej.

Część 2: Heurystyka wyboru wartości.

Przebieg badania:

Dla czterech wybranych sudoku (nr 15,26,42,45) przeprowadziłam badanie składające się z dwóch części:

Część 1:

Dla czterech heurystyk wyboru zmiennej (w kolejności definicji, dynamicznie w kolejności od najmniejszej dziedziny, statycznie rzędami w kolejności najbardziej wypełnionych rzędów oraz statycznie w kolejności najbardziej wypełnionych rzędów, a następnie kolumn) uruchomiłam po dziesięć razy każdą z metod.

Część 2:

Dla czterech heurystyk wyboru wartości (w kolejności definicji, w kolejności losowej dynamicznie, w kolejności najczęściej występujących już na planszy wartości) uruchomiłam po dziesięć razy każdą z metod.

Sudoku nr 15 – problem łatwy

Badanie heurystyk wyboru zmiennej

VarHeu	BTCzas	BTWezly	BTNawroty	BT Czas do 1	BTNawroty do 1	BTWezly do 1
VAR-BY DEFINITION	1,3	9068	2339	0,3	2008	466
VAR-BY DOMAIN DYNAMIC	142,3	839296	216372	11,8	74660	19400
VAR-BY ROWS	1,6	9960	3290	0,1	2778	852
VAR-BY ROWS AND COLUMNS	1,1	7835	2491	0	710	167

Tabela 14 Uśrednione parametry wykonania Backtrackingu dla poszczególnych heurystyk wyboru zmiennej dla Sudoku nr 15.

VarHeu	FCCzas	FCWezly	FC Nawroty	FC Czas do 1	FC Wezly do 1	FC Nawroty do 1
VAR-BY DEFINITION	3	1489	9860	0,1	320	165
VAR-BY DOMAIN DYNAMIC	1,3	803	7140	0	69	8
VAR-BY ROWS	3,6	1601	11680	0,4	441	253
VAR-BY ROWS AND COLUMNS	1,7	1167	9010	0	128	27

Tabela 15 Uśrednione parametry wykonania Forward Checkingu dla poszczególnych heurystyk wyboru zmiennej dla Sudoku nr 15.

Można zauważyć, że dla Backtrackingu wyraźnie najgorsza jest heurystyka w kolejności dziedzin – jest tak dlatego, że wówczas algorytm nie przechodzi kolejno po zależnych od siebie rzędach/kolumnach, a jest to podobne do losowego przechodzenia po zmiennych. Ponieważ dziedziny się nie zmieniają, wszystko zależy od początkowego stanu.

Dla Forward checkingu jest to jednak heurystyka najlepsza, szczególnie w ujęciu dynamicznym – dziedziny są zmienne, stąd algorytm w wyraźny sposób za każdym razem wybiera najlepszą dla siebie zmienną w aktualnym momencie.

W tym przypadku wstępne uporządkowanie danych poprzez ułożenie je w zależności od rzędu i kolumny daje najlepsze rezultaty dla backtrackingu.

Badanie heurystyk wyboru wartości

ValHeu	BTCzas	BTWezly	BTNawroty	BT Czas do 1	BTNawroty do 1	BTWezly do 1
VAL-BY DEFINITION	1,5	9068	2339	0,4	2008	466
VAL-RANDOM	6,1	9068,7	2339	3,9	5847,9	1459,1
VAL - EXISTING VALS	24,1	9070	2339	5,2	1985,1	461,3

Tabela 16 Uśrednione parametry wykonania Backtrackingu dla poszczególnych heurystyk wyboru wartości dla Sudoku nr 15.

ValHeu	FCCzas	FCWezly	FC Nawroty	FC Czas do 1	FC Wezly do 1	FC Nawroty do 1
VAL-BY DEFINITION	2,5	1489	9860	0,1	320	165
VAL-RANDOM	4,3	1489	9860	1,6	723,8	429,4
VAL - EXISTING VALS	4,3	1489	9860	0,5	300	149

Tabela 17 Uśrednione parametry wykonania Forward Checkingu dla poszczególnych heurystyk wyboru wartości dla Sudoku nr 15.

Można zauważyć, że wybór heurystyki wyboru zmiennej nie wpływa w znaczący sposób na liczbę odwiedzonych węzłów oraz nawrotów ogółem. Wpływa jednak na czas działania algorytmu, a także liczbę odwiedzonych węzłów, nawrotów oraz czas do pierwszego rozwiązania, ponieważ decyduje o kolejności rozpatrywania rozwiązań.

Można zauważyć, że heurystyka losowa wypada najgorzej, jeśli chodzi o liczbę węzłów do pierwszego rozwiązania w obydwu algorytmach.

Heurystyka kolejności pod względem wartości już określonych zmiennych wypada lepiej - w obu przypadkach przebywa najmniejszą liczbę węzłów do znalezienia pierwszego rozwiązania. Niestety, jest bardzo kosztowna czasowo, więc zarówno dla Backtrackingu, jak i Forward Checkingu, znalezienie pierwszego rozwiązania jest szybsze dla kolejności z definicji.

Dodatkowo, ma bardzo długi czas wykonania całego algorytmu dla Backtrackingu.

Najbardziej odpowiednią heurystyką jest więc w kolejności definicji, zarówno dla FC jak i BT.

Sudoku nr 26 – problem średni

Badanie heurystyk wyboru zmiennej

VarHeu	BTCzas	BTWezly	BTNawroty	BT Czas do 1	BTWezly do 1	BTNawroty do 1
VAR-BY DEFINITION	151,9	826944	236104	135,7	733156	208640
VAR-BY DOMAIN DYNAMIC	1131	6071807,5	1519614	383,8	2120747,2	531432
VAR-BY ROWS	295,1	1496912	498307	254,1	1288161	427923
VAR-BY ROWS AND COLUMNS	160,7	838094	278289	60,5	312193	107862

Tabela 18 Uśrednione parametry wykonania Backtrackingu dla poszczególnych heurystyk wyboru zmiennej dla Sudoku nr 26.

VarHeu	FCCzas	FCWezly	FC Nawroty	FC Czas do 1	FC Wezly do 1	FCNawroty do 1
VAR-BY DEFINITION	144	93519	614070	127,8	82865	54291
VAR-BY DOMAIN DYNAMIC	5,9	2637	23510	1,1	702	554
VAR-BY ROWS	103,4	62277	449250	89,8	54066	39053
VAR-BY ROWS AND COLUMNS	89,1	52844	406670	33,8	20217	15476

Tabela 19 Uśrednione parametry wykonania Forward Checkingu dla poszczególnych wyboru zmiennej heurystyk dla Sudoku nr 26.

Po raz kolejny da się zauważyć, że wybór zmiennych w kolejności dziedziny nie jest odpowiednią heurystyką dla Backtrackingu – daje znacząco gorsze wyniki. W tym przypadku jest to najlepiej widoczne, ponieważ jest on najbardziej kosztowny obliczeniowo.

Dla Forward Checkingu możemy natomiast zauważyć odmienną tendencję – heurystyka ta pozwala na znalezienie rozwiązania prawie *15 razy szybciej* niż druga najlepsza heurystyka.

Pozostałe heurystyki podobnie, chociaż w tym przypadku dla forwardcheckingu to heurystyka w kolejności rzędów/kolumn i rzędów okazuje się lepsza niż ta z definicji.

Badanie heurystyk wyboru wartości

ValHeu	BTCzas	BTWezly	BTNawroty	BT Czas do 1	BTNawroty do 1	BTWezly do 1
VAL-BY DEFINITION	132,5	826944	236104	117,8	733156	208640
VAL-RANDOM	668,6	826945,4	236104	513,9	635428,1	181289,6
VAL - EXISTING VALS	1988,7	826947	236104	1836,5	763942,6	217760

Tabela 20 Uśrednione parametry wykonania Backtrackingu dla poszczególnych heurystyk wyboru wartości dla Sudoku nr 26.

Tabela 21 Uśrednione parametry wykonania Forward Checkingu dla poszczególnych heurystyk wyboru wartości dla Sudoku nr 26.

ValHeu	FCCzas	FCWezly	FC Nawroty	FC Czas do 1	FC Wezly do 1	FC Nawroty do 1
VAL-BY DEFINITION	125,6	93519	614070	111,1	82865	54291
VAL-RANDOM	266,4	93519	614070	153,4	53757,1	35285,7
VAL - EXISTING VALS	267,1	93519	614070	246,1	85943	56458

Po raz kolejny heurystyka wyboru wartości nie ma większego wpływu na liczbę węzłów oraz nawrotów ogółem. Najlepsza jest heurystyka w kolejności definicji. Gorzej wypada losowa – zarówno czas jest

Sudoku nr 42– problem mieszany 1 (40 rozwiązań)

Badanie heurystyk wyboru zmiennej

VarHeu	BTCzas	BTWezly	BTNawroty	BT Czas do 1	BTNawroty do 1	BTWezly do 1
VAR-BY DEFINITION	74,8	454195	97235	3,9	26042	5482
VAR-BY DOMAIN DYNAMIC	718,2	4420024	1019694	42,7	264775	61421
VAR-BY ROWS	256,4	1429151	404874	14,3	84906	23970
VAR-BY ROWS AND COLUMNS	225,2	1281298	310837	4,8	28152	6831

Tabela 22 Uśrednione parametry wykonania Backtrackingu dla poszczególnych heurystyk wyboru zmiennej dla Sudoku nr 42.

VarHeu	FCCzas	FCWezly	FC Nawroty	FC Czas do 1	FC Wezly do 1	FC Nawroty do 1
VAR-BY DEFINITION	67,9	46808	337890	2,5	2027	1411
VAR-BY DOMAIN DYNAMIC	10,0	5250	47150	0,5	521	404
VAR-BY ROWS	112,6	77644	570790	7,3	5537	4008
VAR-BY ROWS AND COLUMNS	107,2	68549	491110	2,3	1739	1137

Tabela 23 Uśrednione parametry wykonania Forward Checkingu dla poszczególnych heurystyk wyboru zmiennej dla Sudoku nr 42.

Po raz kolejny heurystyka w kolejności zależnej od wielkości dziedziny sprawdza się najlepiej dla Forward Checkingu (sześciokrotnie szybciej niż) oraz najgorzej dla Backtrackingu. Tym razem również heurystyki bazujące na początkowej liczbie wypełnionych pól w kolumnie i rzędzie sprawdzają się znacznie gorzej niż ta z definicji.

Badanie heurystyk wyboru wartości

ValHeu	BTCzas	BTWezly	BTNawroty	BT Czas do 1	BT Wezly do 1	BT Nawroty do 1
VAL-BY DEFINITION	64,5	454195	97235	3	26042	5482
VAL-RANDOM	280,9	454198	97235	32,6	52419,3	11130,5
VAL - EXISTING VALS	1228,8	454325	97235	33	12748	2822

Tabela 24 Uśrednione parametry wykonania Backtrackingu dla poszczególnych heurystyk wyboru wartości dla Sudoku nr 42.

ValHeu	FCCzas	FCWezly	FCNawroty	FC Czas do 1	FC Wezly do 1	FC Nawroty do 1
VAL-BY DEFINITION	59,1	46808	337890	2,1	2027	1411
VAL-RANDOM	127,3	46808	337890	13	4822,7	3400,4
VAL - EXISTING VALS	125,8	46808	337890	4,5	1732	1147

Tabela 25 Uśrednione parametry wykonania Forward Checkingu dla poszczególnych heurystyk wyboru wartości dla Sudoku nr 42.

Wybór heurystyki wyboru wartości nie wpływa na liczbę węzłów oraz nawrotów.

Po raz kolejny najlepiej pod względem czasu sprawdza się heurystyka w kolejności definicji, a heurystyka w zależności od już wpisanych w sudoku wartości pozwala znaleźć pierwsze rozwiązanie na wcześniejszym etapie szukania. Nie jest to jednak efektywne czasowo.

W przypadku Forward Checkingu heurystyka w kolejności już istniejących wartości zmiennych znajduje rozwiązanie nieco szybciej niż losowa, a w przypadku Backtrackingu znacznie wolniej.

Sudoku nr 45– problem mieszany 2 (0 rozwiązań)

Badanie heurystyk wyboru zmiennej

VarHeu	BTCzas	BTWezly	BTNawroty	BT Czas do 1	BTNawroty do 1	BTWezly do 1
VAR-BY DEFINITION	6,9	46265	10913	-	-	-
VAR-BY DOMAIN DYNAMIC	101,1	587607	166868	-	-	-
VAR-BY ROWS	0	31	19	-	-	-
VAR-BY ROWS AND COLUMNS	0	42	23	-	-	-

Tabela 26 Uśrednione parametry wykonania Backtrackingu dla poszczególnych heurystyk wyboru zmiennej dla Sudoku nr 45.

VarHeu	FCCzas	FCNawroty	FCWezly	FC Czas do 1	FC Wezly do 1	FCNawroty do 1
VAR-BY DEFINITION	14,1	8322	65310	-	-	-
VAR-BY DOMAIN DYNAMIC	1	565	5240	-	-	-
VAR-BY ROWS	0	20	180	-	-	-
VAR-BY ROWS AND COLUMNS	0	23	170	-	-	-

Tabela 27 Uśrednione parametry wykonania Forward Checkingu dla poszczególnych heurystyk wyboru zmiennej dla Sudoku nr 45.

W tym przypadku najlepszy wynik dały heurystyki wyboru zmiennej w kolejności rzędów oraz rzędów i kolumn – najprawdopodobniej pozwoliły najszybciej odkryć, że problem nie ma rozwiązania.

Heurystyka w kolejności dziedziny pozwoliła znaleźć rozwiązanie po znacznie dłuższym czasie, dla Forward Checkingu nawet najdłuższym.

Dla heurystyki w kolejności dziedziny w przypadku Fast Checkingu również udało się szybko znaleźć rozwiązanie, a raczej jego brak. Jednak dla Backtrackingu była to najgorsza heurystyka.

Badanie heurystyk wyboru wartości

ValHeu	BT Czas	BT Węzły	BT Nawroty	BT Czas do 1	BT Nawroty do 1	BT Węzły do 1
VAL-BY DEFINITION	6,1	46265	10913	0	0	0
VAL-RANDOM	30,5	46265	10913	0	0	0
VAL - EXISTING VALS	129,4	46265	10913	0	0	0

Tabela 28 Uśrednione parametry wykonania Backtrackingu dla poszczególnych heurystyk wyboru wartości dla Sudoku nr 45.

ValHeu	FC Czas	FC Węzły	FC Nawroty	FC Czas do 1	FC Węzły do 1	FC Nawroty do 1
VAL-BY DEFINITION	11,7	8322	65310	0	0	0
VAL-RANDOM	25,5	8322	65310	0	0	0
VAL - EXISTING VALS	23,2	8322	65310	0	0	0

Tabela 29 Uśrednione parametry wykonania Forward Checkingu dla poszczególnych heurystyk wyboru wartości dla Sudoku nr 45.

W przypadku problemu bez rozwiązań trudno mówić o jakichkolwiek różnicach pomiędzy heurystykami wyboru wartości – nie ma mowy o parametrach znalezienia pierwszego rozwiązania, a z pozostałych jedyną różnicę stanowi czas.

Ten, podobnie jak przy poprzednich problemach, dla Backtrackingu jest najkrótszy dla w kolejności definicji, a najdłuższy dla heurystyki w zależności od istniejących wartości. Dla Forward Checkingu również najkrótszy dla w kolejności definicji, ale najdłuższy dla losowych.

Wnioski:

Badanie heurystyk wyboru zmiennej

Dla **Backtrackingu** najlepszą heurystyką okazała się ta **w kolejności definicji**, a nieco gorzej (choć często niewiele) spisywały się heurystyki w kolejności rzędów oraz kolumn i rzędów. Stąd można zauważyć, że ważniejsze od ułożenia początkowego jest to, aby kolejne zmienne były od siebie zależne. **Przy heurystyce w kolejności definicji wypełnia się nie tylko kolejne rzędy i kolumny, ale także kwadraty** – dlatego kiedy wykorzystujemy kolejność rzędu i kolumny często tracimy tę przewagę. Pomimo że zyskujemy pracę na rzędach i kolumnach, które są lepsze, to będąc oddalone od siebie, rzadziej pozwalają na szybkie odrzucenie wyniku nieodpowiedniego w zakresie kwadratu.

Heurystyka w kolejności wielkości dziedziny nie sprawdza się dla Backtrackingu z tego samego powodu – powoduje, że sprawdzane są zmienne oddalone od siebie (tu już nie tylko w obrębie kwadratów, ale też rzędów i kolumn). Ponieważ dziedziny są stałe w tym algorytmie, kolejność zawsze faworyzuje jedynie początkowe ułożenie planszy. (Warto dodatkowo zauważyć, że jako że algorytm nie wykonuje za każdym razem sortowania zmiennych pod względem dziedziny, więc długi czas nie jest z tym związany)

Dla **Forward Checkingu** zdecydowanie najlepszą heurystyką jest ta **w kolejności dziedziny**, ponieważ dynamicznie sprawdza ona zmiany w najważniejszym elemencie tego algorytmu.

Pozostałe heurystyki działają bardzo podobnie do Backtrackingu – najczęściej najlepsza jest heurystyka w kolejności definicji, a następnie w kolejności rzędów oraz w kolejności rzędów i kolumn, choć wszystkie te heurystyki dają podobne wyniki.

Badanie heurystyk wyboru wartości

Heurystyki wyboru wartości nie mają większego wpływu na liczbę węzłów i nawrotów ogółem, a jedynie na czas znalezienia pierwszej wartości.

Badanie wykazało, że najlepszą pod względem czasu jest heurystyka wyboru wartości w kolejności definicji, zarówno dla Forward Checkingu jak i Backtrackingu.

Można jednak zauważyć, że heurystyka pod względem liczby istniejących już przypisanych wartości, wcześniej (jeśli chodzi o numer węzła) znajduje rozwiązanie, ale szczególnie w przypadku Backtrackingu zajmuje znacznie więcej czasu. Świadczy to jednak o pewnej skuteczności metody, stąd być może przy zmianie implementacji metoda ta mogłaby okazać się najlepsza.

Najgorszą heurystyką jest losowa – znajduje pierwsze rozwiązanie znacznie później niż pozostałe dwa algorytmy. Jest wprawdzie szybsza dla Backtrackingu niż druga z heurystyk, ale nie jest bardziej skuteczna. Jest tak dlatego, że w pozostałych przypadkach kolejne sprawdzane wartości są sobie bliższe, natomiast w kolejności losowej są one niezależne od siebie.