Value Compression of Pattern Databases

Nathan R. Sturtevant, *University of Denver*; Ariel Felner, *BGU*; Malte Helmert, *University of Basel*

Pattern Database Compression

What is a Pattern Database (PDB)?

A PDB is a heuristic that estimates the distance to the goal for search algorithms such as A*. It computes and stores the distances in an abstract state space and stores them in a table.

Entry Compression (Felner et al, 2007) compresses the PDB by combining entries and storing the minimum. Thus, there are fewer entries in the compressed PDB.

Entry Compression (EC)

original PDB

1	3	4	8	9	2	3	10

1	4	2	3

2x compressed PDB

The minimum value is stored so that the heuristic is still admissible (non-overestimating).

Value Compression (VC) [new]

VC keeps all entries in the PDB, but reduces the number of bits per entry. This reduces the ranges of values that can be stored.

original PDB (≥4 bits)

1	3	4	8	9	2	3	10

1	3	4	7	7	2	3	7

compressed PDB (3 bits)

Value Compression: No wasted bits

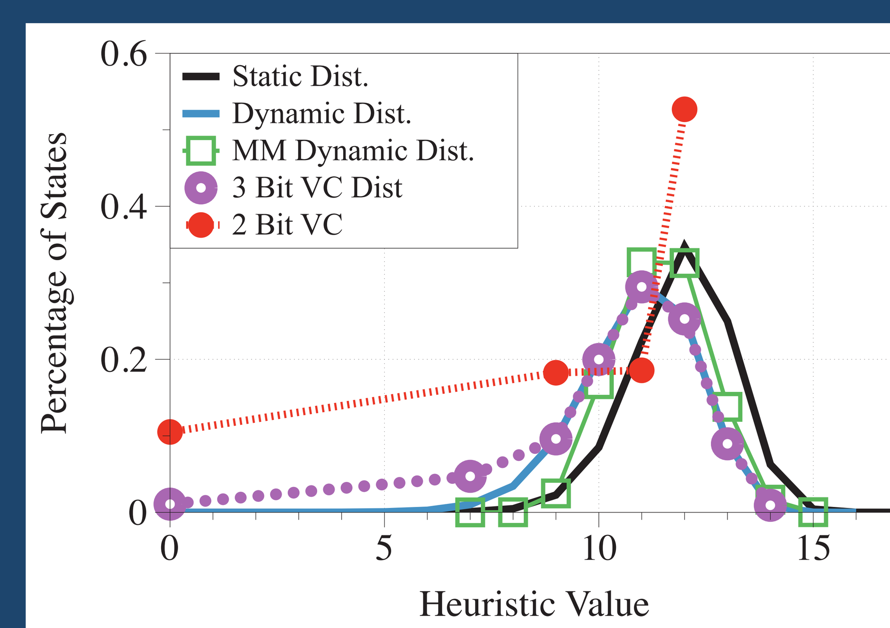
Value Compression Details

- Value Compression (VC) divides the heuristic values into ranges $R_1 \ldots R_n$. Ranges are stored instead of values.
- The minimum value of the range is used during search to maintain admissibility.
- Dynamic programming is used to optimize the ranges and maximize the average h-value in the compressed PDB.

Example of EC vs VC in Top Spin

	1,760MB	880MB			440MB	
D	Total	$\underline{VC2}$	$VC2_{\bar{h}}$	EC2	$VC4_{\bar{h}}$	EC4
0	1	1	12	2	10,188,753	4
1	11	11		22		40
2	94	94	94	186		340
3	731	731	731	1,430		2,596
4	5,353	5,353	5,353	10,340		18,736
5	37,275	37,275	37,275	70,894		127,756
6	245,468	245,468	245,468	457,304		813,700
7	1,508,099	1,508,099	1,508,099	2,722,458		4,724,408
8	8,391,721	8,391,721	8,391,721	14,408,820		23,870,392
9	40,012,497	40,012,497	40,012,497	63,502,746	190,013,262	97,318,252
10	150,000,765	150,000,765	150,000,765	212,692,340		290,434,356
11	393,482,172	393,482,172	393,482,172	478,114,034	393,482,172	553,276,900
12	612,084,904	612,084,904	612,084,904	601,419,722	1,170,638,373	549,750,508
13	440,655,534	440,655,534	440,655,534	328,304,534		217,340,348
14	110,437,757	110,437,757	110,437,757	59,883,892		26,009,144
15	7,389,524	7,460,178	7,389,524	2,721,910		634,464
16	**70,633**		70,654	11,924		616
17	**21**			2		
Avg.	11.90	11.90	**11.90**	11.59	**11.38**	11.27

Runtime Heuristic Distribution



Experimental Results

General Observations

- Conventional wisdom says the low values in the PDB are more important than the high values. So, removing low values from the PDB should hurt performance.
- In practice, the most common heuristic values in the runtime distribution (looked up during search) must be preserved.
- Bidirectional Pathmax (BPMX) is crucial for local propagation of heuristic values and recovery of lost information.
- VC is most effective when the number of values in the PDB is just larger than the nearest power of two.
- EC can be effectively combined with VC.

Experimental Setup

The paper contains experiments on many domains and algorithms. We report on 18-4 Top Spin here.

We compare compression factors in IDA* with BPMX as well as combinations of EC and VC. Combinations with VC have the best performance (bold).

Memory	EC	VC	VC-bits	Nodes	Time
1	1	1	8	3.88M	15.29
0.5 (A)	1	2	4	**3.88M**	**15.32**
0.375	1	2.66	3	4.03M	15.44
0.25 (B)	1	4	2	10.39M	33.63
0.5 (A)	2	1	8	7.11M	27.70
0.25 (B)	2	2	4	**7.11M**	**27.88**
0.1875	2	2.66	3	7.37M	28.44
0.125 (C)	2	4	2	30.43M	80.04
0.25 (B)	4	1	8	13.75M	51.06
0.125 (C)	4	2	4	**13.74M**	**50.97**
0.094	4	2.66	3	14.31M	51.52
0.0625	4	4	2	30.48M	77.68

Bidirectional Search: MM

- The MM algorithm guarantees that the search frontiers meet in the middle.
- Small heuristic values aren't used.

Memory	EC	VC	VC-bits	Nodes	Time
1	1	1	8	2.07M	30.60
0.5	2	1	8	3.55M	55.61
0.5	1	2	4	**2.07M**	**30.93**
0.25	4	1	8	5.42M	83.87
0.25	2	2	4	**3.55M**	**55.55**
0.25	1	4	2	4.63M	66.46
0.125	4	2	4	5.42M	84.06
0.125	2	4	2	**5.19M**	**79.17**

Moving AI Lab
NSF Award 1551406
UNIVERSITY of DENVER
DANIEL FELIX RITCHIE SCHOOL OF ENGINEERING & COMPUTER SCIENCE
Ben-Gurion University of the Negev