A Measure of Quality for IDA* Heuristics

Robert Clausecker Florian Schintke

Zuse Institute Berlin

The 14th Annual Symposium on Combinatorial Search
July 27, 2021

Iterative Deepening A*

Single-source shortest path algorithm due to Korf (1985)

- find shortest path $v_0, v_1, \ldots, v_n$ from source vertex v_0 to goal vertex $z = v_n$ by depth first search
- a maximum depth n is fixed and iteratively increased until a path is found
- search assisted by a *heuristic function* $h(v)$ like with A*
- $h(v)$ gives a *lower bound* for the distance from v to z
- if $h(v)$ plus the number of steps already taken to get to v exceeds n, v is pruned from the search tree

Fundamental Questions

Fundamental Questions

- how many nodes will one IDA* search expand?

Fundamental Questions

- how many nodes will one IDA* search expand?
- how does $h(v)$ influence the number of expanded nodes?

Fundamental Questions

- how many nodes will one IDA* search expand?
- how does $h(v)$ influence the number of expanded nodes?
- how can we tune the heuristic to prune more nodes?

The KRE formula

Analysis of IDA*'s time complexity due to Korf, Reid, and Edelkamp (2001).

- one iteration of IDA* is expected to expand $E(v_0, d, P)$ nodes

$$E(v_0, d, P) = \sum_{i=0}^{d} N_i P(d - i) \tag{1}$$

- N_i: number of vertices at distance i from v_0
- $P(i) = \mathrm{P}[\, h(v) \leq i\,]$ for equilibrium distributed vertex v

The KRE formula

Analysis of IDA*'s time complexity due to Korf, Reid, and Edelkamp (2001).

- one iteration of IDA* is expected to expand $E(v_0, d, P)$ nodes

$$E(v_0, d, P) = \sum_{i=0}^{d} N_i P(d - i) \tag{1}$$

- N_i: number of vertices at distance i from v_0
- $P(i) = \mathrm{P}[\, h(v) \leq i \,]$ for equilibrium distributed vertex v
- precise modelling, but beware the assumptions made (cf. Zahavi et al. 2010)

Limitations

The KRE formula...

Limitations

The KRE formula...

- depends on the neighborhood N_i of v_0

Limitations

The KRE formula...

- depends on the neighborhood N_i of v_0
- is linked to the search depth

Limitations

The KRE formula...

- depends on the neighborhood N_i of v_0
- is linked to the search depth
- makes the influence of $h(v)$ hard to quantify

Simplifying the KRE formula

$$E(v_0, d, P) = \sum_{i=0}^{d} N_i P(d-i)$$

Simplifying the KRE formula

$$E(v_0, d, P) = \sum_{i=0}^{d} N_i \, P(d-i)$$

- use the *branching factor* b to model N_i as b^i

$$E(b, d, P) = \sum_{i=0}^{d} b^i \, P(d-i) \qquad (2)$$

Simplifying the KRE formula

$$E(v_0, d, P) = \sum_{i=0}^{d} N_i \, P(d-i)$$

- use the *branching factor* b to model N_i as b^i

$$E(b, d, P) = \sum_{i=0}^{d} b^i \, P(d-i) \tag{2}$$

- untangle the contribution of N_i and d from that of $P(d-i)$

$$E(b, d, P) = b^d \sum_{i=0}^{d} P(i) / b^i \tag{3}$$

Simplifying the KRE formula

$$E(v_0, d, P) = \sum_{i=0}^{d} N_i P(d-i)$$

- use the *branching factor* b to model N_i as b^i

$$E(b, d, P) = \sum_{i=0}^{d} b^i P(d-i) \tag{2}$$

- untangle the contribution of N_i and d from that of $P(d-i)$

$$E(b, d, P) = b^d \sum_{i=0}^{d} P(i)/b^i \tag{3}$$

- make the result largely independent of d

$$E(b, d, P) \approx b^d \sum_{i=0}^{\infty} P(i)/b^i \tag{4}$$

The Heuristic Quality η

Using $p(i) = \mathrm{P}[\, h(v) = i\,]$.

$$\eta = \sum_{i=0}^{\infty} \frac{p(i)}{b^i} = \frac{b-1}{b} \sum_{i=0}^{\infty} \frac{P(i)}{b^i} \tag{5}$$

The Heuristic Quality η

Using $p(i) = \mathrm{P}[\, h(v) = i\,]$.

$$\eta = \sum_{i=0}^{\infty} \frac{p(i)}{b^i} = \frac{b-1}{b} \sum_{i=0}^{\infty} \frac{P(i)}{b^i} \tag{5}$$

$$E(b, d, P) \approx \frac{b}{b-1} b^d \eta \tag{6}$$

The Heuristic Quality η

Using $p(i) = \mathrm{P}[\, h(v) = i\,]$.

$$\eta = \sum_{i=0}^{\infty} \frac{p(i)}{b^i} = \frac{b-1}{b} \sum_{i=0}^{\infty} \frac{P(i)}{b^i} \tag{5}$$

$$E(b, d, P) \approx \frac{b}{b-1} b^d \eta \tag{6}$$

$$\frac{b}{b-1} b^d \eta - \frac{1}{b-1} \leq E(b, d, P) < \frac{b}{b-1} b^d \eta \tag{7}$$

The Heuristic Quality η

Using $p(i) = \mathrm{P}[\, h(v) = i\,]$.

$$\eta = \sum_{i=0}^{\infty} \frac{p(i)}{b^i} = \frac{b-1}{b} \sum_{i=0}^{\infty} \frac{P(i)}{b^i} \tag{5}$$

$$E(b, d, P) \approx \frac{b}{b-1} b^d \eta \tag{6}$$

$$\frac{b}{b-1} b^d \eta - \frac{1}{b-1} \leq E(b, d, P) < \frac{b}{b-1} b^d \eta \tag{7}$$

$\longrightarrow$ uninformed search has $\eta = 1$

Computing the Quality

How do we compute η for a given heuristic?

Computing the Quality

How do we compute η for a given heuristic?

- compute $p(i)$ for all i and plug the values into the definition

Computing the Quality

How do we compute η for a given heuristic?

- compute $p(i)$ for all i and plug the values into the definition
- for pattern databases: tally up PDB entries

Computing the Quality

How do we compute η for a given heuristic?

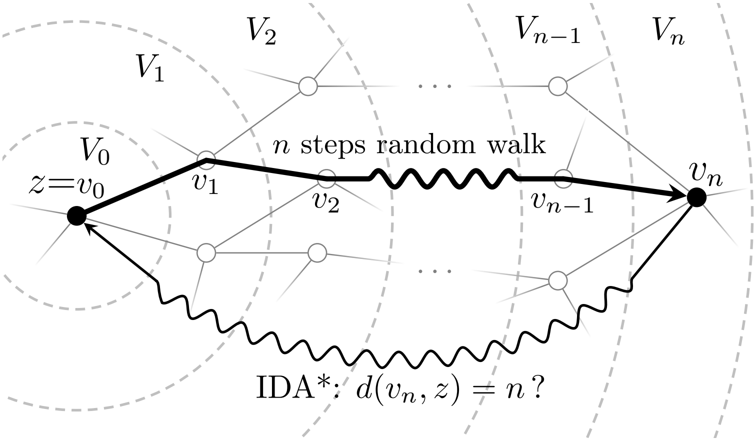
- compute $p(i)$ for all i and plug the values into the definition
- for pattern databases: tally up PDB entries
- in general: sample $h(v)$ for random vertices and estimate the distribution
- not quite as simple as it looks (random uniform sampling does not suffice)

Sphere Stratified Sampling

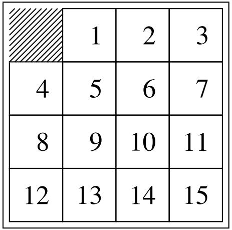
Strategy:

- split the search space into *spheres* of vertices equidistant to z
- sample vertices from each sphere individually
- $h(v)$ values correlate strongly with $d(v, z)$
- lower sampling error $\longrightarrow$ less samples are needed

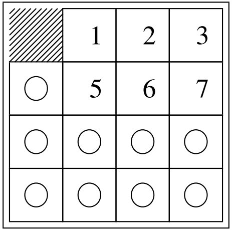
Sphere Stratified Sampling



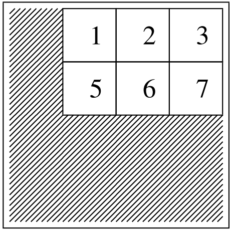
Sliding Tile Puzzles



15 puzzle

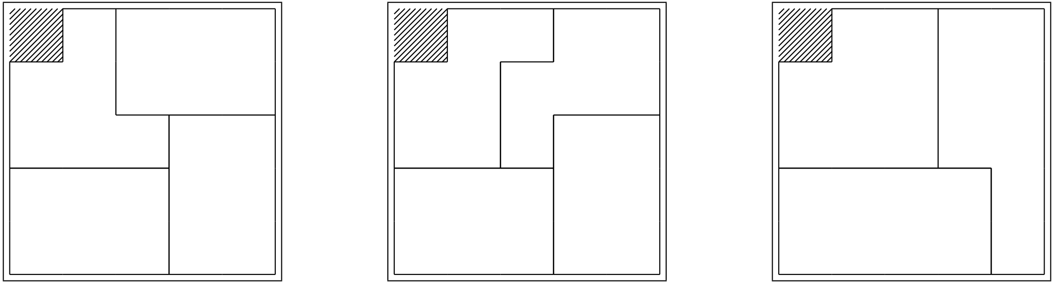


non-additive PDB



additive PDB

PDB Partitionings



(a) (b) (c)

Qualities of some 24 Puzzle Heuristics

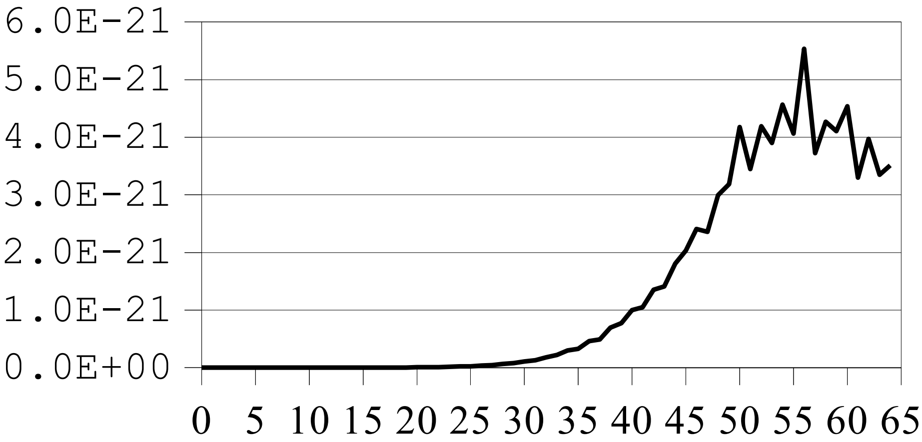
heuristic		*quality*	*error*
Manhattan heuristic		$\eta = 9.926 \times 10^{-20}$	(9.08 %)
partitioning (a)	w/o transposition search	$\eta = 4.562 \times 10^{-22}$	(0.73 %)
partitioning (a)	with transposition search	$\eta = 1.359 \times 10^{-22}$	(0.39 %)
partitioning (b)	w/o transposition search	$\eta = 4.391 \times 10^{-22}$	(2.13 %)
partitioning (b)	with transposition search	$\eta = 1.611 \times 10^{-22}$	(0.57 %)
partitioning (c)	w/o transposition search	$\eta = 1.097 \times 10^{-22}$	(0.49 %)
partitioning (c)	with transposition search	$\eta = 6.780 \times 10^{-23}$	(0.33 %)
small collection	w/o transposition search	$\eta = 8.548 \times 10^{-23}$	(0.50 %)
small collection	with transposition search	$\eta = 3.787 \times 10^{-23}$	(0.35 %)
large collection	w/o transposition search	$\eta = 6.751 \times 10^{-23}$	(0.71 %)
large collection	with transposition search	$\eta = 3.208 \times 10^{-23}$	(0.24 %)

Quality Histograms

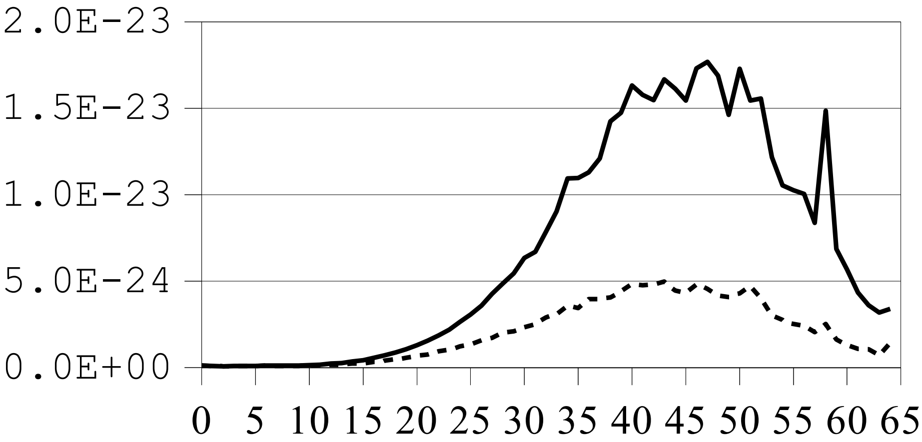
What parts of the search space are most relevant to search performance?

- plot contribution η_k to η by sphere V_k
- higher contribution $\longrightarrow$ more time is spent at vertices of this distance
- can infer most important region of h values from most important spheres

Quality Histograms: Manhattan Heuristic



Quality Histograms: Partitioning (a)



Quality Histograms: Small Collection