

funções heurísticas de procura

a situação

problemas difíceis

- formulação difícil

- espaço de procura de dimensão muito grande

a função heurística pode ajudar muito

- se for bem escolhida

dimensão do espaço de estados

ex: o puzzle de deslizar de 3x3

- espaço de estados da procura

fator de ramificação (médio) $2,67 \sim 3$

vazio: 2 movimentos nos cantos

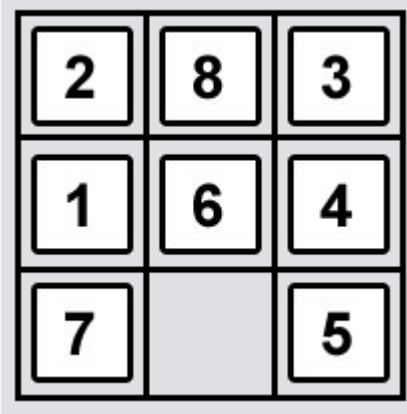
3 movimentos nos lados

4 movimentos no centro

média de movimentos até à solução ~ 22

expansão exaustiva da árvore: $\sim 3^{22} \sim 3.1 \times 10^{10}$

em grafo (estados únicos possíveis): $9!/2 = 181.440$



2	8	3
1	6	4
7		5

só metade das
configurações
de um *array*
abstrato de 3x3

dimensões grandes do EE

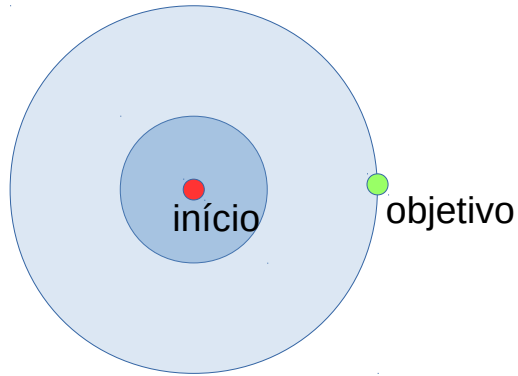
puzzle de 4x4

1	2	8	11
9	3	6	14
4	7	12	10
5	15	13	

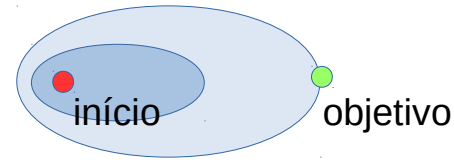
em grafo (estados únicos possíveis): $16!/2 \sim 10^{13}$

crescimento exponencial com n de EE de puzzle $n \times n$

A* vs. custo uniforme (recapitulando)



custo uniforme: $f(n) = g(n)$



a importância
da função
heurística

A*: $f(n) = g(n) + h(n)$

admissível:

≥ 0

nunca sobrestima o
custo até ao objetivo

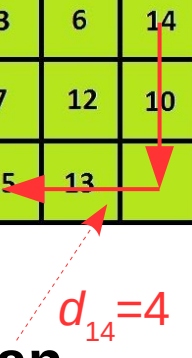
heurística do puzzle $n \times n$

- h_1 = número de peças mal colocadas

neste caso $h_1 = 13$

admissível: $h_1 \geq 0$ e otimista!

1	2	8	11
9	3	6	14
4	7	12	10
5	15	13	



- h_2 = soma das distâncias das peças à posição final

distância de Manhattan

neste caso $h_2 = 0+0+2+3+1+2+1+4+5+2+1+2+2+1+2 = 28$

1 2 8 11 9 3 6 14 4 7 12 10 5 15 13

admissível: $h_2 \geq 0$ e

menor ou igual ao nº de movimentos necessários

qualidade da heurística

- uma medida da qualidade é
fator de ramificação efetivo b^*

se A^* gera N nós na procura

e a profundidade da solução é d

então (contando com a raiz) a árvore completa de prof. d

$$N + 1 = 1 + b^* + (b^*)^2 + \dots + (b^*)^d$$

ex: se A^* encontra a solução a prof. 5 com 52 nós, $b^* = 1,9$

uma boa heurística
tem $b^* \sim 1 + \Delta$
(próximo de 1)

d	Search Cost (nodes generated)			Effective Branching Factor		
	BFS	$A^*(h_1)$	$A^*(h_2)$	BFS	$A^*(h_1)$	$A^*(h_2)$
6	128	24	19	2.01	1.42	1.34
8	368	48	31	1.91	1.40	1.30
10	1033	116	48	1.85	1.43	1.27
12	2672	279	84	1.80	1.45	1.28
14	6783	678	174	1.77	1.47	1.31
16	17270	1683	364	1.74	1.48	1.32
18	41558	4102	751	1.72	1.49	1.34
20	91493	9905	1318	1.69	1.50	1.34
22	175921	22955	2548	1.66	1.50	1.34
24	290082	53039	5733	1.62	1.50	1.36
26	395355	110372	10080	1.58	1.50	1.35
28	463234	202565	22055	1.53	1.49	1.36

Figure 3.26 Comparison of the search costs and effective branching factors for 8-puzzle problems using breadth-first search, A^* with h_1 (misplaced tiles), and A^* with h_2 (Manhattan distance). Data are averaged over 100 puzzles for each solution length d from 6 to 28.

fonte: AIMA 4ª ed.

comparação

- A^* muito melhor do que procura em largura (BFS)
- h_2 melhor do que h_1

porquê?

$h_2(n) \geq h_1(n)$, para qualquer nó n h_2 domina h_1

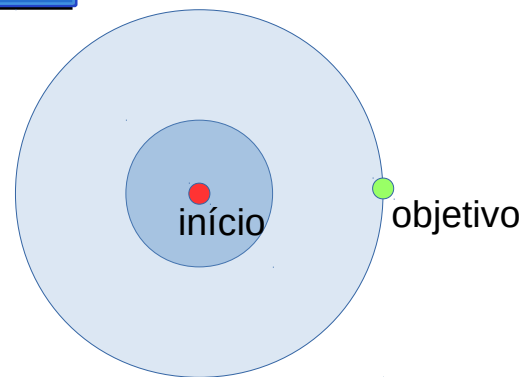
qualquer nó com $f(n) < C^*$ é expandido

i.e. com $h(n) < C^* - g(n)$

$\Rightarrow h_1(n)$ expande pelo menos tantos nós como $h_2(n)$
e em muitos casos ainda mais alguns

heurísticas limite

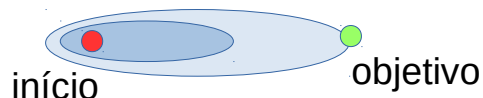
se $h(s) = 0$, então A^* é idêntico a
procura de custo uniforme



se $h(s) = \text{CustoFuturoMínimo}(s)$, então A^* só expande nós
num caminho de custo mínimo



em geral $h(s)$ está entre estes dois extremos

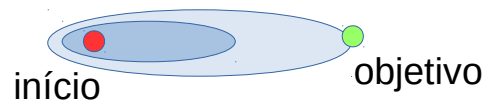


compromissos

uma heurística dominante é uma melhor estimativa do custo até ao objetivo

mais bem dirigida

expande menos nós



hélas!

normalmente é mais demorada de calcular

=> avaliar se compensa o ganho na expansão de nós...

como obter boas heurísticas?

basta relaxar...



geração de heurísticas

frequentemente, heurísticas admissíveis correspondem a soluções de problemas relaxados

- com menos restrições

- com ações mais grosseiras

e pode até ser útil começar por uma heurística inadmissível e “repará-la”

relaxar o puzzle $n \times n$

ações no puzzle real

uma peça pode mover-se da casa A para a casa B se
A for adjacente, vertical ou horizontalmente, a B e B vazia

relaxando cada uma e as duas:

- a) uma peça pode mover-se da casa A para a casa B se A e B adjacentes
- b) uma peça pode mover-se da casa A para a casa B se B vazia
- c) uma peça pode mover-se da casa A para a casa B

heurística h_1

heurística h_2

heurística de problema relaxado

- custo exato do problema relaxado tem de verificar a desigualdade triangular, logo usar esse custo como heurística do problema real produz uma heurística **consistente**
- problema relaxado adiciona arcos ao grafo do EE

uma solução ótima do problema original é também solução do problema relaxado

mas o problema relaxado pode ter melhores soluções

logo, custo da solução exata do problema relaxado é uma heurística **admissível** do problema original (é sempre \leq custo real)

e se?...

tivermos várias heurísticas admissíveis mas nenhuma domina qualquer das outras? qual escolher?

uma que as usa a todas!

$$h(n) = \max \{h_1(n), \dots, h_m(n)\}$$

é admissível, consistente e domina todas as componentes

heurísticas a partir de subproblemas

- podem obter-se heurísticas admissíveis a partir do custo da solução de um subproblema

ex: resolver a colocação de
1, 2, 3 e 4

é claramente um limite inferior
da solução do problema original
em alguns casos é mais exata
do que a dist. Manhattan

*	2	4
*		*
*	3	1

	1	2
3	4	*
*	*	*

e gerar bases de dados de padrões

com soluções exatas para subproblemas
que devem depois ser adicionáveis
(padrões disjuntos)

subproblemas disjuntos

solução de 1-2-3-4 deve mover 5-6-7-8

não são problemas disjuntos

logo, a soma das heurísticas não é uma heurística admissível

*	2	4
*		*
*	3	1

	1	2
3	4	*
*	*	*

mas se só contarmos os movimentos de 1, 2, 3 e 4 e de 5, 6, 7 e 8, nos respetivos subproblemas

a soma dos dois custos é um limite inferior do custo do problema original (heurística admissível)

bases de dados de padrões disjuntos (só em alguns probls.)

aprendizagem de heurísticas

um agente ao resolver um problema obtém o custo real até ao objetivo desde cada estado usado na solução

dados de várias resoluções podem permitir aprender $h(n)$

torna-se a tarefa mais fácil se forem definidas propriedades relevantes dos estados, úteis para a heurística

ex: nº de peças mal colocadas no puzzle

agente pode vir a inferir que custo médio de 5 peças mal colocadas é, p.ex. 14

heurística admissível pode ser (p.ex.) dividir por 2 a média do custo

exercício

- um agricultor tem de atravessar um rio com um lobo, uma cabra e uma couve
- o bote só pode levar o agricultor e mais 2 deles
- o lobo não pode ficar só com a cabra
- a cabra não pode ficar só com a couve



atravessar o rio

- uma heurística
- outra heurística?
- comparar A^* com PL, PF, PCU
- e considerar pedir uma redefinição do problema!...

