

# Adapting Novelty to Classical Planning as Heuristic Search

李梓桐 卢彦作 贺刘丁

SUN YAT-SEN UNIVERSITY

2020 年 11 月 29 日

- 1 Background: 背景知识回顾
- 2 Novelty Heuristics for Classical Planning: 算法核心介绍
- 3 Experimental Evaluation: 实验结果评估

# Background

经典的规划问题 (Classical Planning) 由哪几部分组成?

- 经典的规划任务  $\Pi = \langle \mathcal{V}, \mathcal{O}, s_0, s_* \rangle$

- 经典的规划任务  $\Pi = \langle \mathcal{V}, \mathcal{O}, s_0, s_* \rangle$
- $\mathcal{V}$ : 一组有限域状态的变量
  - $\mathcal{D}(v)$ : 每个变量  $v \in \mathcal{V}$  与其关联的有限域 (维度)

- 经典的规划任务  $\Pi = \langle \mathcal{V}, \mathcal{O}, s_0, s_* \rangle$
- $\mathcal{V}$ : 一组有限域状态的变量
  - $\mathcal{D}(v)$ : 每个变量  $v \in \mathcal{V}$  与其关联的有限域 (维度)
- $\mathcal{O}$ : 一组有限的操作, 每个操作由  $\langle pre, eff \rangle$ , 即  $\mathcal{V}$  中的一对元素组成, 称为前提和结果

- 经典的规划任务  $\Pi = \langle \mathcal{V}, \mathcal{O}, s_0, s_* \rangle$
- $\mathcal{V}$ : 一组有限域状态的变量
  - $\mathcal{D}(v)$ : 每个变量  $v \in \mathcal{V}$  与其关联的有限域 (维度)
- $\mathcal{O}$ : 一组有限的操作, 每个操作由  $\langle pre, eff \rangle$ , 即  $\mathcal{V}$  中的一对元素组成, 称为前提和结果
- $s_0$ :  $\Pi$  的起始状态



- 经典的规划任务  $\Pi = \langle \mathcal{V}, \mathcal{O}, s_0, s_* \rangle$
- $\mathcal{V}$ : 一组有限域状态的变量
  - $\mathcal{D}(v)$ : 每个变量  $v \in \mathcal{V}$  与其关联的有限域 (维度)
- $\mathcal{O}$ : 一组有限的操作, 每个操作由  $\langle pre, eff \rangle$ , 即  $\mathcal{V}$  中的一对元素组成, 称为前提和结果
- $s_0$ :  $\Pi$  的起始状态
- $s_*$ :  $\Pi$  的目标状态

- state: 对  $\mathcal{V}$  的一个完整赋值
  - $\mathcal{S} = \prod_{v \in \mathcal{V}} D(v)$  是  $\Pi$  的状态空间
  - 将操作  $o$  应用于状态  $s$ , 用  $s[o]$  表示
  - 对于一个变量  $v \in \mathcal{V}$  和一个状态  $s$ ,  $v$  在  $s$  中的值用  $s[v]$  表示

- state: 对  $\mathcal{V}$  的一个完整赋值
  - $\mathcal{S} = \prod_{v \in \mathcal{V}} D(v)$  是  $\Pi$  的状态空间
  - 将操作  $o$  应用于状态  $s$ , 用  $s[o]$  表示
  - 对于一个变量  $v \in \mathcal{V}$  和一个状态  $s$ ,  $v$  在  $s$  中的值用  $s[v]$  表示
- fact: 对于一个变量  $v \in \mathcal{V}$  和一个值  $\theta \in D(v)$ , 由其组成的对  $\langle v, \theta \rangle$  称为一个 fact。fact 是 state 的部分, 多个 fact 组成一个 state
  - $\mathbb{F}$ : 所有  $\Pi$  中 fact 的集合

Definition 1 (Reward Novelty) 给出一个奖励函数  $R: S \mapsto \mathbb{R}^{0+}$  和到目前为止看到的一组历史状态  $S$ , 一个 fact (variable value) 的新颖性 (novelty) 得分  $f$  由以下定义:

$$N(f, S, R) = \begin{cases} \max_{s \in S, f \in s} R(s), & f \in s \text{ for some } s \in S \\ -\infty, & \text{otherwise.} \end{cases}$$

$$N(f, S, R) = \begin{cases} \max_{s \in S, f \in s} R(s), & f \in s \text{ for some } s \in S \\ -\infty, & \text{otherwise.} \end{cases}$$

简单来说, 对于一个新状态中的 fact,  $N(f, S, R)$  就是在历史状态中挑选出包含该 fact 的状态, 返回这些状态中的最大奖励值。

$$N(f, S, R) = \begin{cases} \max_{s \in S, f \in s} R(s), & f \in s \text{ for some } s \in S \\ -\infty, & \text{otherwise.} \end{cases}$$

若历史记录中不存在包含该 fact 的状态，也就是说这个 fact 从未出现过，则返回负无穷。如果这个新状态  $s$ ，当  $R(s) > N(f, S, R)$  时，我们称这个新状态是 novel 的。

即一个 state  $s$  被认为是 novel 的, iff  $R(s) > N(f, S, R)$  for some  $f \in s$ .

# Novelty Heuristics for Classical Planning

# Classical Planning

For now we restrict our attention to the deterministic case.

- Complete initial state specifications
- Deterministic effects of actions

What we will cover

- Representing and reasoning about actions in the situation calculus
- Planning in STRIPS
- Planning as search



Definition 2 (Heuristic Novelty) Given a heuristic function  $h : \mathcal{S} \mapsto \mathbb{R}^{0+}$  and a set of states seen so far  $S$ , the novelty score of a fact (variable value)  $f$  is defined as

$$N(f, S, h) = \begin{cases} \min_{s \in S, f \in s} h(s), & f \in s \text{ for some } s \in S \\ \infty, & \text{otherwise.} \end{cases}$$

Given a state  $s$ , the novelty score of a fact  $f$  in state  $s$  is defined as  $N(f, s, S, h) = N(f, S, h) - h(s)$  if  $f \in s$

$$N(f, S, h) = \begin{cases} \min_{s \in S, f \in s} h(s), & f \in s \text{ for some } s \in S \\ \infty, & \text{otherwise.} \end{cases}$$

$$N(f, s, S, h) = N(f, S, h) - h(s)$$

- a fact is novel in state  $s$  if its novelty score in  $s$  is strictly positive
- a state is novel if it contains at least one novel fact

$$h_{\text{BN}}(s) = \begin{cases} 0, & \exists f \in s, N(f, s, S, h) > 0 \\ 1 & \text{otherwise} \end{cases}$$

$$h_{\text{BN}}(s) = \begin{cases} 0, & \exists f \in s, N(f, s, S, h) > 0 \\ 1 & \text{otherwise} \end{cases}$$

只要新状态存在某个赋值是 novel 的，启发式函数值就是 0，否则为 1。

$$h_{\text{QN}}(s) = |\mathcal{V}| - \sum_{f \in s} N^+(f, s, S, h)$$

- $N^+(f, s, S, h)$  is 1 when  $N(f, s, S, h) > 0$  and 0 otherwise
- Similarly, let  $N^-(f, s, S, h)$  be 1 when  $N(f, s, S, h) < 0$  and 0 otherwise

$$h_{QN}(s) = |\mathcal{V}| - \sum_{f \in s} N^+(f, s, S, h)$$

- $N^+(f, s, S, h)$  is 1 when  $N(f, s, S, h) > 0$  and 0 otherwise
- Similarly, let  $N^-(f, s, S, h)$  be 1 when  $N(f, s, S, h) < 0$  and 0 otherwise

进一步地,  $h_{QN}$  将赋值为 novel 的数量考虑进来, 赋值为 novel 的越多, 启发式函数值越小。

quantified both

$$h_{QB}(s) = \begin{cases} h_{QN}(s), & h_{QN}(s) < |\mathcal{V}| \\ |\mathcal{V}| + \sum_{f \in s} N^-(f, s, S, h), & \text{otherwise} \end{cases}$$

$$h_{QB}(s) = \begin{cases} h_{QN}(s), & h_{QN}(s) < |\mathcal{V}| \\ |\mathcal{V}| + \sum_{f \in s} N^-(f, s, S, h), & \text{otherwise} \end{cases}$$

更进一步,  $h_{QB}$  将赋值不 novel 的也考虑进来, 赋值不 novel 的越多, 启发式函数值越大。



$$h_{\text{QF}}(s) = \begin{cases} k|\mathcal{V}| - N_k^+(s), & N_k^+(s) > 0 \\ k|\mathcal{V}| + N_k^-(s), & \text{otherwise} \end{cases}$$

- $N_k^+(f, s) = \begin{cases} k, & N(f, s) = \infty \\ \left\lceil \frac{k \times N(f, s)}{M_s} \right\rceil, & 0 < N(f, s) < \infty \\ 0, & \text{otherwise} \end{cases}$
- $N_k^-(f, s) = \begin{cases} -\left\lceil \frac{k \times N(f, s)}{M_s} \right\rceil, & N(f, s) < 0 \\ 0, & \text{otherwise} \end{cases}$
- $N_k^{\text{op}}(s) = \sum_{f \in \mathcal{S}} N_k^{\text{op}}(f, s)$ , where  $\text{op} \in \{+, -\}$

# Multiple Base Estimators

Definition 3 Given a set of states seen so far  $S$ , a set of heuristic functions  $H = \{h_1, \dots, h_n \mid h_i : \mathcal{S} \mapsto \mathbb{R}^{0+}\}$ , and a state  $s$ , the novelty of a fact  $f$  in state  $s$  is defined as

$$N(f, s, S, H) = \max_{h \in H} N(f, s, S, h) \quad \text{if} \quad f \in s$$

# Experimental Evaluation

# Experimental Evaluation

Coverage		$h^{FF}$	$h_{BN}^{FF}$	$h_{QN}^{FF}$	$h_{QB}^{FF}$	$h_{QF}^{FF}$	$h^{LM}$	$h_{BN}^{LM}$	$h_{QN}^{LM}$	$h_{QB}^{LM}$
airport	50	31	39	38	39	<b>50</b>	29	32	32	37
barman11	20	4	14	14	17	9	<b>20</b>	13	13	13
depot	22	15	18	20	<b>21</b>	20	14	18	19	19
driverlog	20	18	<b>20</b>	<b>20</b>	<b>20</b>	<b>20</b>	18	18	18	19
elevators08	30	11	10	10	9	11	23	<b>30</b>	<b>30</b>	27
elevators11	20	0	0	0	0	0	5	<b>10</b>	<b>10</b>	<b>10</b>
floortile11	20	<b>8</b>	7	6	6	<b>8</b>	0	0	0	0
freecell	80	<b>80</b>	79	<b>80</b>	79	79	<b>80</b>	<b>80</b>	<b>80</b>	<b>80</b>
Sum	<b>1143</b>	729	906	911	<b>924</b>	825	790	845	842	852
Sum total	<b>1456</b>	1042	1219	1224	<b>1237</b>	1138	1103	1158	1155	1165

Table 1: Coverage for GBFS, comparing novelty variants  $h_{BN}$ ,  $h_{QN}$ ,  $h_{QB}$ , and  $h_{QF}$  for  $h$ , tie breaking on  $h$  to the base heuristic  $h$  for  $h \in \{h^{FF}, h^{LM}\}$ .

这一部分对于使用不同的 novelty variants 的结果，总体上可见使用 novelty variants 要更优

# Experimental Evaluation

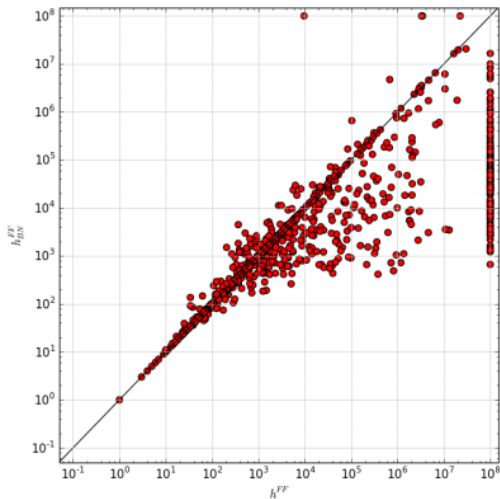
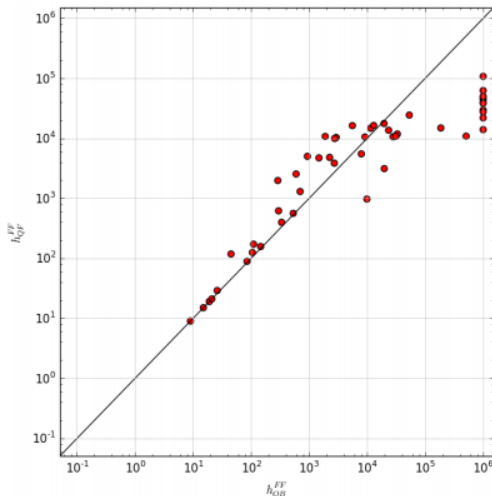


图: 计算点数的对比

# Experimental Evaluation

原文还比较了几种不同的 novelty variants 的效果, 得到了在总体上  $h_{QN}^{FF}$  略优于  $h_{BN}^{FF}$ ,  $h_{QB}^{FF}$  优于  $h_{QB}^{FF}$  以及  $h_{QF}$  的表现最差



# Experimental Evaluation

Coverage	FF LM	* FF LM			LM FF	* LM FF		
		$h_{QB}^{FF}$	$h_{QB}^{LM}$	$h_{QB}^{FF,LM}$		$h_{QB}^{LM}$	$h_{QB}^{FF}$	$h_{QB}^{FF,LM}$
airport (50)	35	37	<b>42</b>	40	30	38	37	35
barman11 (20)	3	17	14	10	17	<b>20</b>	<b>20</b>	18
depot (22)	16	19	18	<b>21</b>	16	18	20	<b>21</b>
driverlog (20)	17	<b>20</b>	18	<b>20</b>	15	19	<b>20</b>	<b>20</b>
elevators08 (30)	10	13	29	29	13	29	12	<b>30</b>
elevators11 (20)	0	0	9	9	0	<b>10</b>	0	<b>10</b>
floortile11 (20)	4	<b>6</b>	4	5	3	4	5	3
freecell (80)	78	<b>80</b>	79	78	79	79	<b>80</b>	78
grid (5)	<b>5</b>	<b>5</b>	4	<b>5</b>	4	4	<b>5</b>	<b>5</b>
logistics98 (35)	22	26	20	<b>27</b>	19	20	<b>27</b>	<b>27</b>
mprime (35)	31	<b>35</b>	32	<b>35</b>	31	32	<b>35</b>	<b>35</b>
mystery (30)	18	18	18	<b>19</b>	17	18	18	18
<b>Sum (1033)</b>	663	851	828	886	754	836	867	<b>895</b>
<b>Sum total (1456)</b>	1086	1274	1251	1309	1177	1259	1290	<b>1318</b>

这一部分是把不同的启发式函数结合起来的实验结果

第三部分的实验则是引入了另外两种变体  $h_{RB}$ , 与之前使用的  $h_{FF}$  与  $h_{LM}$  选取的队列交替进行

Coverage	$h^{RB}$	$h_{QB}^{RB}$	$[h^{RB}, h^{LM}]$	$[h_{QB}^{RB}, h^{LM}]$
airport (50)	34	39	32	<b>43</b>
barman11 (20)	18	<b>20</b>	19	<b>20</b>
depot (22)	19	<b>20</b>	18	<b>20</b>
driverlog (20)	<b>20</b>	<b>20</b>	19	<b>20</b>
floortile11 (20)	<b>7</b>	<b>7</b>	4	3
freecell (80)	<b>80</b>	79	78	<b>80</b>
grid (5)	4	<b>5</b>	<b>5</b>	<b>5</b>
mprime (35)	<b>35</b>	<b>35</b>	31	<b>35</b>
mystery (30)	18	<b>19</b>	18	<b>19</b>
nomystery11 (20)	14	18	15	<b>19</b>
<b>Sum (722)</b>	591	625	548	<b>634</b>
<b>Sum total (1456)</b>	1325	1359	1282	<b>1368</b>

可见结合使用都比原来的性能要更优



- 针对已有的启发式估计提出了一种 Novelty 的概念
- 提出了一种利用这一概念进行规划的方法
- 展示了与其他基本启发式函数结合应用的结果，发现其能够改善性能

- 其中的某些方法在一些特定的 domain 中表现较好，问题可能是由于这些算法是 problem specific 的，需要研究更加通用的方法
- 本文对于 Multiple Base Estimators 的计算过于简单，应该可以找到更好的办法进行估计
- 本文提出的关于状态的 novelty 是独立的，而且这个定义利用了所有的单个事实而不是事实集，实际上这个定义应该也可以用于事实集

Thank you!