

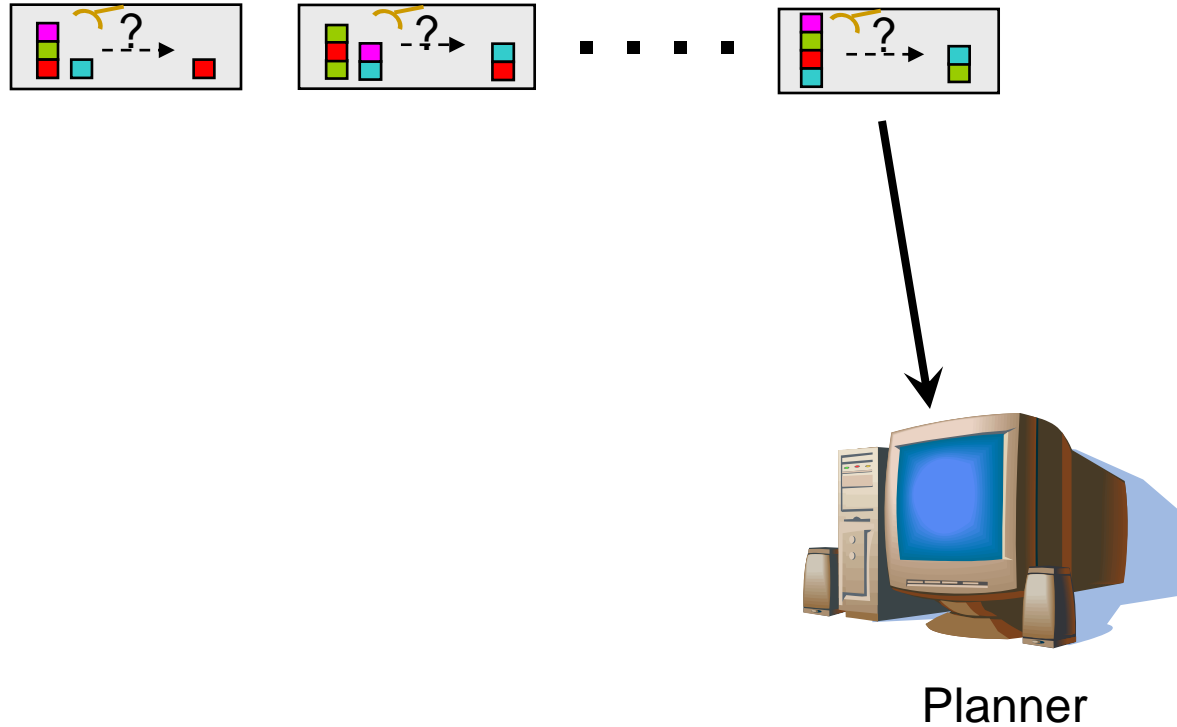
---

# Results of IPC 2008: **Learning Track**

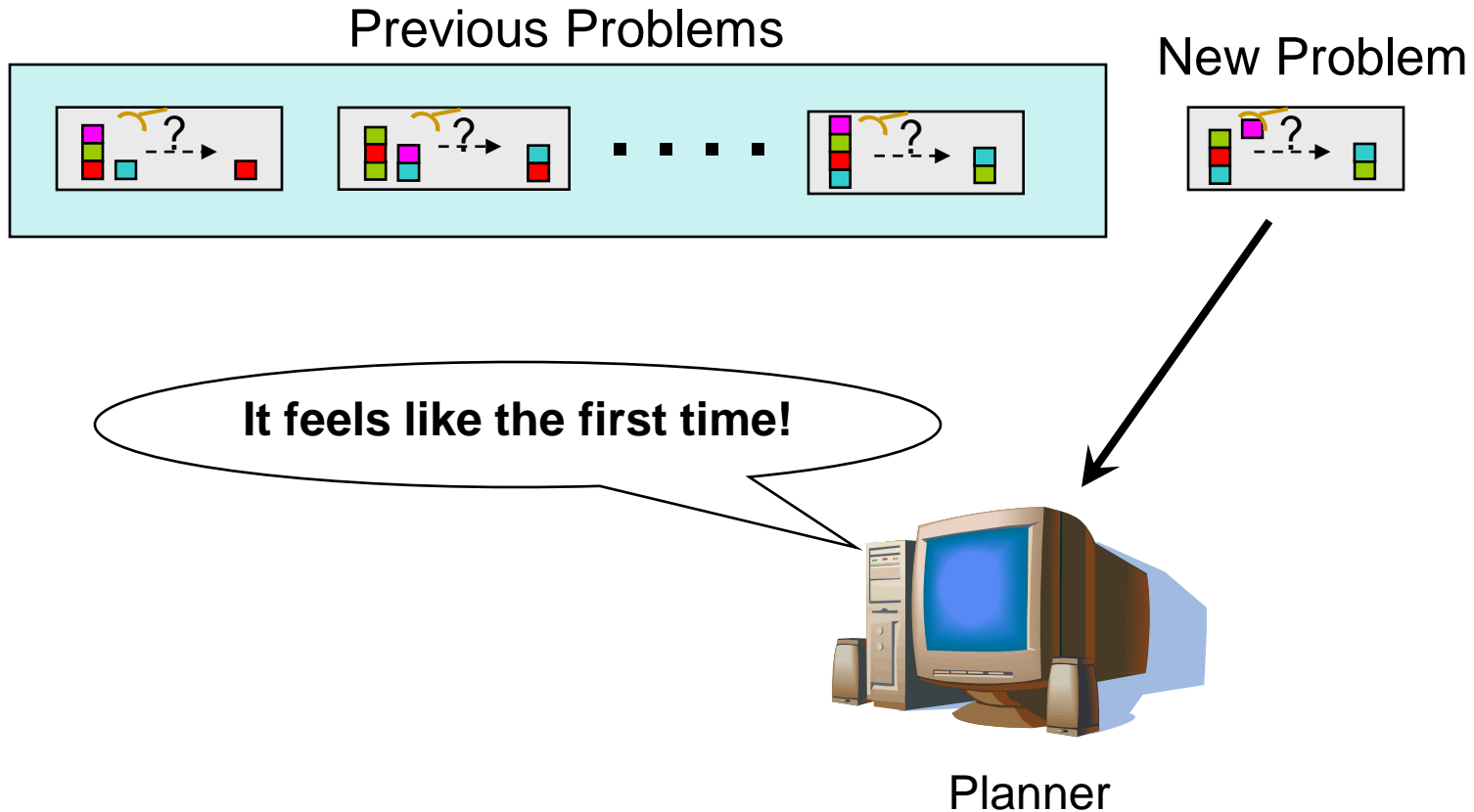
---

**Organizers:** Alan Fern, Prasad Tadepalli, Roni Khardon

# Motivation

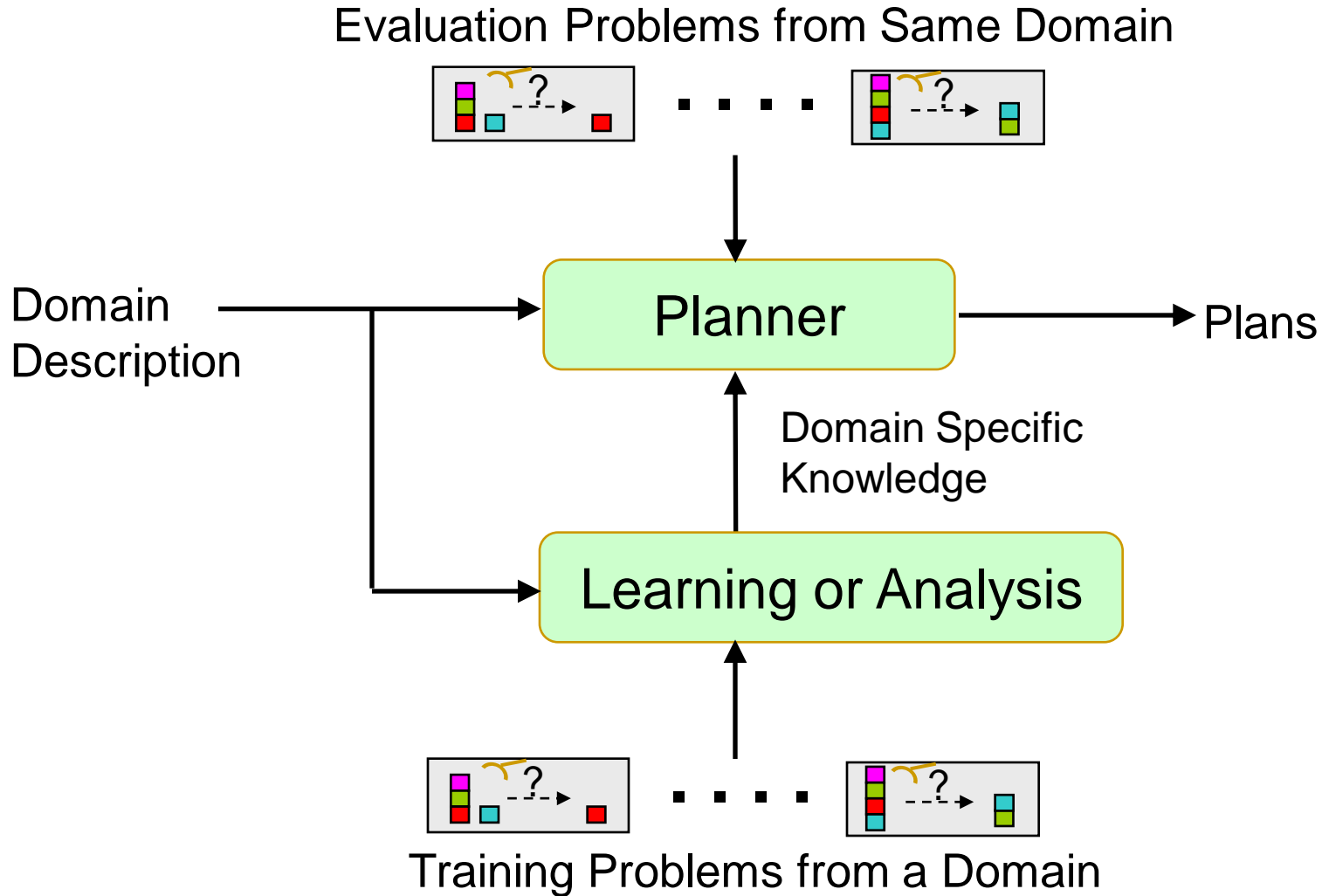


# Motivation



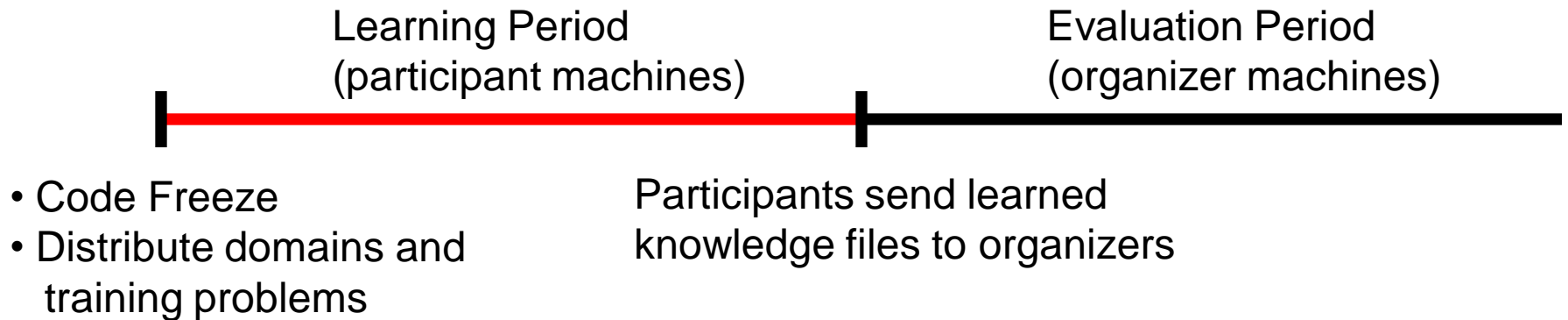
- AI systems should **learn** from past experience
- Learning track designed to encourage work on learning for planning

# Learning Domain Specific Knowledge



# Competition Format

1. Code freeze of learning and planning components
2. Domains and training problems released
3. Learning systems given 2 week learning period run on participant machines
4. Each team sends the organizers the resulting learned knowledge files (one per domain)
5. Evaluation period conducted on organizer machines



---

# Language and Objective Functions

- STRIPS Fragment of PDDL
  - Planning Time Metric
    - A planner scores 0 for unsolved problems
    - Otherwise, score for a problem is  $T^*/T$ 
      - $T$  = time for planner to solve problem
      - $T^*$  = minimum time for any planner to solve problem
    - A planners overall score is sum over all problems
  - Planning Quality Metric
    - Similar, except score for solved problems is  $N^*/N$ 
      - $N$  = plan length,  $N^*$  = min length over all planners
-

---

# Domains (details on web-site)

- GoldMiner
    - Use different types of bombs in a grid-style mine to uncover gold
  - Matching Blocksworld
    - Blocksworld with polarized blocks and hands, must use hand of correct polarity
  - Classic N-Puzzle
    - Traditional problem studied in macro learning literature
  - Parking
    - Rearrange parked cars on a grounded street
  - Sokoban
    - Simple instances of classic computer puzzle game
  - Thoughtful Solitaire
    - A solitaire variant
-

---

# Training and Evaluation Problems

- Provided 60 training problems per domain divided into two sets
    - **30 bootstrap problems** : small, easy problems
    - **30 target problems** : harder problems representative of evaluation problems
  - Evaluation used 30 problems per domain of the harder variety
    - 15 minute time limit per problem
  - All problem sets, domain definitions, and problem generators available on the web site
-



---

# Participants

- Great turnout!
    - 13 systems from 10 different groups
  - Abstracts of the systems will be on the web-page
  - The systems can be roughly categorized into 4 types based on the type of knowledge learned and how it is used:
    - Learn **policies** that are incorporated into search for plans
    - Learn **macro actions** sequences
    - Learn **sub-goal decomposition** knowledge
    - Learn **value functions** that serve as search heuristics
    - Learn to **configure a portfolio** of planners
-

# Participants: Policy Learners

- **CABALA** (Tomas de la Rosa; Daniel Borrajo; Angel Garcia Olaya)
  - Learns case base to guide a look-ahead heuristic planner
- **Roller** (Tomas de la Rosa; Sergio Jimenez)
  - Learns decision-tree policies to guide look-ahead planner
- **REPLICA** (Rocio Garcia-Duran; Fernando Fernandez; Daniel Borrajo)
  - Learns instance-based policies to guide enforced hill-climbing
- **ObtuseWedge** (Sungwook Yoon)
  - Learns reactive policies that are incorporated into FF's BFS
- **Sayphi-Rules** (Susana Fernandez Arregui; Daniel Borrajo)
  - Learns decision tree policies to guide enforced hill-climbing

---

# Participants: Macro Learners

- Wizard+FF and Wizard+SGPlan (M.A. Hakim Newton; John Levine; Maria Fox; Derek Long)
  - Learns macro actions to be used by a base planner
  
- Macro-AltAlt (Murugeswari I; N. S. Narayanaswamy)
  - Learns macro actions that are used by the planner AltAlt



---

# Participants: Sub-problem Decomposition

- **DEA1 and DEA2** (Jacques Bibai; Perre Savant; Marc Schoenauer; Vincent Vidal)
  - Learns knowledge to search over sub-goal decompositions which are solved via CPT



---

# Participants: Portfolio Configuration

- **PbP.s** (Beniamino Galvani; Alfonso E. Gerevini; Alessandro Saetti; Mauro Vallati)
  - Learns knowledge to configure a portfolio of domain-independent planners including Fast Downward, Metric-FF, LPG-td, MacroFF, Marvin, SGPLAN5, YAHSP



---

# Participants: Value Function Learners

- Relational Function Approximation 1 & 2 (RFA1,RFA2)  
(Jia-Hong Wu; Robert Givan)
  - Learn linear value functions are used as FF's heuristic



# Results: Overall Time Metric

System	Time Metric Score (max 180)	% Solved
<b>PbP.s</b>	<b>80.2</b>	93
ObtuseWedge	76.2	65
Wizard-SGP	59.3	51
Wizard-FF	45.3	57
Macro-AltAlt	15.8	42
RFA1	11.7	48
Roller	8.9	31
Replica	7.6	32
Sayphi-Rules	4.0	26
RFA2	2.6	26
DEA2	0.01	18
DEA1	0.01	18
Cabala	0.001	2

Score over all  
180 problems

---

# Time Metric Winner

PbP.s

**Beniamino Galvani, Alfonso Gerevini,  
Alessandro Saetti, Mauro Vallati**

University of Brescia

---



# Per-domain Time Metric Winners

Domain	Winner	Time Metric (max 30)
Gold Miner	Wizard-FF	24.4
Matching BW	PbP.s	25.8
N-Puzzle	ObtuseWedge	29.3
Parking	ObtuseWedge	28.1
Sokoban	Wizard-SGP	27.0
Thoughtful	PbP.s	23.0

Each domain has a dominant planner that achieves close to the maximum score.

# Results: Overall Quality Metric

System	Quality Metric Score (max 180)	% Solved
<b>PbP.s</b>	<b>126.7</b>	93
ObtuseWedge	95.1	65
Wizard-FF	91.2	57
Wizard-SGP	81.1	51
Macro-AltAlt	70.9	42
RFA1	63.9	48
RFA2	35.7	26
DAE1	31.7	18
DAE2	31.4	18
Sayphi-Rules	29.3	26
Replica	25.0	32
Roller	25.0	31
Cabala	2.6	2

---

# Quality Metric Winner

PbP.s

**Beniamino Galvani, Alfonso Gerevini,  
Alessandro Saetti, Mauro Vallati**

University of Brescia

---

# Per-domain Quality Metric Winners

Domain	Winner	Quality Metric (max 30)
Gold Miner	DAE1	28.7
Matching BW	Macro-AltAlt	24.1
N-Puzzle	ObtuseWedge	24.5
Parking	ObtuseWedge	25.5
Sokoban	Wizard-SGP	29.4
Thoughtful	RFA1	19.2

---

# Best Learner?

- Congratulations to PbP.s developers for being the overall winner of both metrics!
  - But, how much are these systems benefitting from learning?
  - We evaluated each system both with and without the learned knowledge
    - Results in a set of 26 experiments (2 per system)
    - We ranked all of these systems using our metrics
-

# Time Metric (top 10)

System	Time Metric Score (max 180)	% Solved
PbP.s (no learning)	74.9	92
<b>PbP.s</b>	69.7	93
<b>ObtuseWedge</b>	61.4	65
<b>Wizard-SGP</b>	46.7	51
<b>Wizard-FF</b>	41.6	57
Wizard-SGP (no learning)	28.8	53
ObtuseWedge (no learning)	25.3	48
Wizard-FF (no learning)	20.2	63
Sayphi-Rules (no learning)	18.0	24
<b>Macro-AltAlt</b>	12.4	42

**Best performer is PbP.s w/o knowledge!**

# Quality Metric (top 10)

System	Quality Metric Score (max 180)	% Solved
<b>PbP.s</b>	121.0	93
PbP.s (no learning)	120.8	92
Wizard-FF (no learning)	106.3	63
<b>ObtuseWedge</b>	91.2	65
<b>Wizard-FF</b>	89.0	57
Wizard-SGP (no learning)	83.3	53
<b>Wizard-SGP</b>	79.6	51
Macro-AltAlt (no learning)	73.9	46
<b>Macro-AltAlt</b>	67.1	42
<b>RFA1</b>	62.3	48

**Best performer is PbP.s with knowledge!**

---

# Analysis

- We see that PbP.s performs about the same overall without learning
    - The performance is largely due to using a very good planner rather than learning!
  - Which system demonstrated the best learning ability?
    - Difficult question to answer
    - Can't simply look at percent improvement after learning over all systems
      - E.g. going from 0 to 0.1 gives infinite improvement
  - We focused on top 5 learning systems to select a best learner
-



# Best Learner Award

- Look at delta for each metric: “Score after learning” – “Score before learning”
- Positive when learning improves metric
- Negative when learning hurts metric

Top 5 Learning System	Delta Time Score	Delta Quality Score	Delta % Solved
PbP.s	- 5.26	0.20	1
ObtuseWedge	<b>36.05</b>	<b>29.02</b>	<b>17</b>
Wizard-FF	21.42	- 17.33	- 6
Wizard-SGP	17.90	- 3.63	- 2
Macro-AltAlt	1.16	- 6.80	- 4

---

# Quality Metric Winner

## Obtuse Wedge

**Sungwook Yoon**

**Palo Alto Research Center**

---

---

# Comparison with Non-Learning Track

- To help judge the performance of systems in our track we compared to planners from non-learning track
  - Ran 11 planners from sequential, satisficing track on learning track problems
    - Thanks to Malte Helmert for running these trials!
    - Did not compare times since run on different machines
    - Compared using quality metric
-

# Inter-Track Comparison(top 10)

System	Quality Metric Score (max 180)	Success Rate	Track
Lama (no learning)	<b>132.63</b>	0.79	Sequential
<b>PbP.s</b>	110.5	<b>0.93</b>	Learning
PbP.s (no learning)	109.8	0.92	Learning
Wizard-FF (no learning)	98.1	0.63	Learning
<b>Wizard-FF</b>	82.2	0.57	Learning
<b>ObtuseWedge</b>	80.4	0.65	Learning
Wizard-SGP (no learning)	78.1	0.53	Learning
<b>Wizard-SGP</b>	75.4	0.51	Learning
Macro-AltAlt (no learning)	66.1	0.46	Learning
FFHA (no learning)	63.73	0.37	Sequential

Our evaluation problem sets are clearly non-trivial with respect to difficulty

---

# Conclusions

- We had great participation (13 systems)
    - **Best Time and Quality Award**: PbP.s
    - **Best Learner Award**: ObtuseWedge
  
  - Clear evidence of successful and significant learning
    - First time such a blind evaluation has been carried out for learning-based planners
  
  - The most successful learning systems leveraged code from state-of-the-art planners
    - Is probably important to be competitive in the future
-

---

# Conclusions

- Learning **did not** improve overall performance compared to best non-learning planners
    - PbP did about the same before and after learning
    - The Lama system outperformed all learning systems except for PbP.s
  - Robustness is a major issue
    - Learning can often make performance degrade
  - Thus, the goal of reliably outperforming non-learning planners via learning is still very much open
-

# IPC-08: Learning Track Results

## Time Metric Results

Systems Domains	Cabala	Dae1	Dae2	Macro AltAlt	ObtuseWedge	PbP.s	Replica	Roller	Savahi Rules	Wizard-FF	Wizard-SGP	RFA1	RFA2
GoldMiner	0.00	0.01	0.01	6.58	9.38	4.42	5.14	6.37	3.85	<b>24.40</b>	23.25	0.00	0.29
Matching BW	0.00	0.00	0.00	5.28	2.03	<b>25.85</b>	0.00	0.00	0.00	0.52	0.00	0.02	0.00
N-Puzzle	0.00	0.00	0.00	3.91	<b>29.33</b>	7.10	0.00	0.26	0.09	2.76	4.42	0.63	0.00
Parking	0.00	0.00	0.00	0.00	<b>28.08</b>	8.96	2.42	2.27	0.00	0.85	0.00	0.70	0.56
Sokoban	0.00	0.00	0.00	0.00	4.42	10.82	0.00	0.00	0.01	8.48	<b>26.99</b>	0.17	0.14
Thoughtful	0.00	0.00	0.00	0.00	3.42	<b>23.02</b>	0.00	0.00	0.00	8.24	4.62	10.21	1.58
<b>Overall</b>	0.001	0.01	0.01	15.77	76.65	<b>80.16</b>	7.56	8.91	3.95	45.25	59.29	11.73	2.57

**Overall Time Metric Winner: PbP.s**

## Quality Metric Results

Systems Domains	Cabala	Dae1	Dae2	Macro AltAlt	ObtuseWedge	PbP.s	Replica	Roller	Savahi Rules	Wizard-FF	Wizard-SGP	RFA1	RFA2
GoldMiner	0.00	<b>28.69</b>	28.41	27.65	17.46	23.96	7.97	7.85	16.14	25.80	25.01	7.88	13.45
Matching BW	1.89	2.00	2.94	<b>24.12</b>	5.69	20.22	1.59	1.10	0.00	12.70	0.59	4.63	0.00
N-Puzzle	0.00	0.00	0.00	19.16	<b>24.50</b>	17.77	0.38	0.45	8.78	13.17	9.82	11.25	0.00
Parking	0.74	0.00	0.00	0.00	<b>25.54</b>	19.48	14.22	15.61	0.00	8.67	0.00	10.60	7.95
Sokoban	0.00	0.97	0.00	0.00	15.27	27.24	0.88	0.00	4.40	17.11	<b>29.40</b>	10.30	4.12
Thoughtful	0.00	0.00	0.00	0.00	6.62	18.01	0.00	0.00	0.00	13.76	16.30	<b>19.28</b>	10.19
<b>Overall</b>	2.63	31.66	31.35	70.92	95.07	<b>126.68</b>	25.05	25.02	29.32	91.21	81.12	63.93	35.70

**Overall Quality Metric Winner: PbP.s**

## Success Rate (fraction of problems solved)

Systems Domains	Cabala	Dae1	Dae2	Macro AltAlt	ObtuseWedge	PbP.s	Replica	Roller	Savahi Rules	Wizard-FF	Wizard-SGP	RFA1	RFA2
GoldMiner	0.00	<b>1.00</b>	<b>1.00</b>	0.97	0.67	<b>1.00</b>	<b>1.00</b>	<b>1.00</b>	<b>1.00</b>	<b>1.00</b>	<b>1.00</b>	0.30	0.47
Matching BW	0.07	0.07	0.10	0.87	0.27	<b>0.93</b>	0.10	0.07	0.00	0.50	0.03	0.33	0.00
N-Puzzle	0.03	0.00	0.00	0.00	<b>1.00</b>	0.87	0.70	0.60	0.00	0.30	0.00	0.53	0.40
Parking	0.00	0.03	0.00	0.00	0.70	<b>1.00</b>	0.03	0.00	0.20	0.63	1.00	0.43	0.30
Sokoban	0.00	0.03	0.00	0.00	0.70	<b>1.00</b>	0.03	0.00	0.20	0.63	1.00	0.43	0.30
Thoughtful	0.00	0.00	0.00	0.00	0.27	<b>0.97</b>	0.00	0.00	0.00	0.47	0.57	0.67	0.40
<b>Overall</b>	0.02	0.18	0.18	0.42	0.65	<b>0.93</b>	0.32	0.31	0.26	0.57	0.51	0.47	0.26

**Best Success Rate: PbP.s**

Comparison of systems' overall performance before and after learning (top 8 shown)

Systems with learning

System	Time Metric	Quality Metric	Success Rate
PbP.s (no learning)	<b>74.9</b>	120.8	0.92
<b>PbP.s</b>	69.7	<b>121.0</b>	<b>0.93</b>
ObtuseWedge	61.4	91.2	0.65
Wizard-SGP	46.7	79.6	0.51
Wizard-FF	41.6	89.0	0.57
Wizard-SGP (no learning)	28.8	83.3	0.53
ObtuseWedge (no learning)	25.3	62.2	0.48
Wizard-FF (no learning)	20.2	106.3	0.63

Compare the difference in performance before and after learning for top 5 (positive values means learning helped)

System	Delta Time Metric	Delta Quality Metric	Delta Success Rate
PbP.s	- 5.26	0.20	1
ObtuseWedge	<b>36.05</b>	<b>29.02</b>	<b>17</b>
Wizard-FF	21.42	- 17.33	- 6
Wizard-SGP	17.90	- 3.63	- 2
Macro-AltAlt	1.16	- 6.80	- 4

**Best Learner Award: ObtuseWedge**