Learning Predictive Models to Configure Planning Portfolios

Isabel Cenamor
Tomás de la Rosa
Fernando Fernández

June 11, 2013

Universidad Carlos III de Madrid

[plg]

Outline

Motivation

- The best planners for each domain in sequential satisfying track (IPC-2011)

Domain	Planner
barman	fd-autotune-1
elevators	forkuniform
floortile	fd-autotune-2
nomystery	arvand
openstacks	fd-autotune-2
parcprinter	arvand
parking	lama-2011

Domain	Planner
pegsol	lama-2011
scanalyzer	arvand
sokoban	fd-autotune-1
tidybot	lamar
transport	roamer
visitall	daeyahsp
woodworking	fdss1

- The portfolios of planners are interesting because there is not a best planner for all domains

Definition

Intuition: "Assign the available time to a sub-set of the available planners, and run this configuration"

The Definition of Planner Portfolio we will use

Given a set of base planners, $\{pl_1, \ldots, pl_n\}$, and a maximum execution time, T, a planning portfolio can be considered as a sequence of m pairs $< pl_1, t_1 >, \ldots, < pl_m, t_m >$, where $pl_i \in \{pl_1, \ldots, pl_n\}$ and $\sum_{j=1}^{m} t_j = T$.

Portfolio Approaches

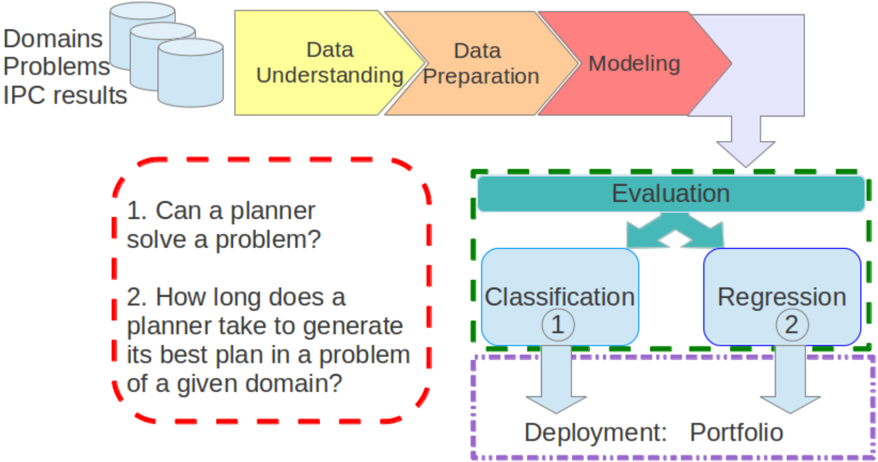
The portfolio configuration can be done:

- Over all seen benchmarks: unique configuration
- Per domain: same configuration per domains
- Per problem: different configuration per instance

Portfolio Approaches	Configuration
FD-Stone-Soup [Fawcett et al., 2011]	Over all
PbP [Gerevini et al., 2009, Gerevini et al., 2011]	Domain
Our approach	Instance

Objectives

- Configure a planning portfolio using predictive models
- Learn these models in function of the objective
  - Whether a planner will be able to find a solution
  - How long it will take
- Create some strategies to combine the models
- Study different evaluation procedures


Domains
Problems
IPC results
Data Understanding
Data Preparation
Modeling
1. Can a planner solve a problem?
2. How long does a planner take to generate its best plan in a problem of a given domain?
Evaluation
Classification
1
Regression
2
Deployment: Portfolio

We created some features to characterize the problem

1. Some of them extracted from the PDDL files (3):
    1. Number of goals
    2. Number of literals
    3. Number of objects
2. A set of elaborated features generated from the problem translation to the SAS+ formalism (41)
    1. Numbers of nodes in the causal graph
    2. Ratio between weights and number of edges
    3. etc.
3. And the class is the solution of the problems (IPC 2011 solutions of Sequential Satisfying Track)(2)
    1. Quality
    2. Time

Data Modeling

We use some algorithms from Weka [Hall et al., 2009] machine learning toolkit to train models and make predictions

1. For classification:

   Decision trees J48
   Instance-based learning algorithms (KNN) [IBk $k = 1, 3, 5$]
   Support vector machine SMO

2. For regression:

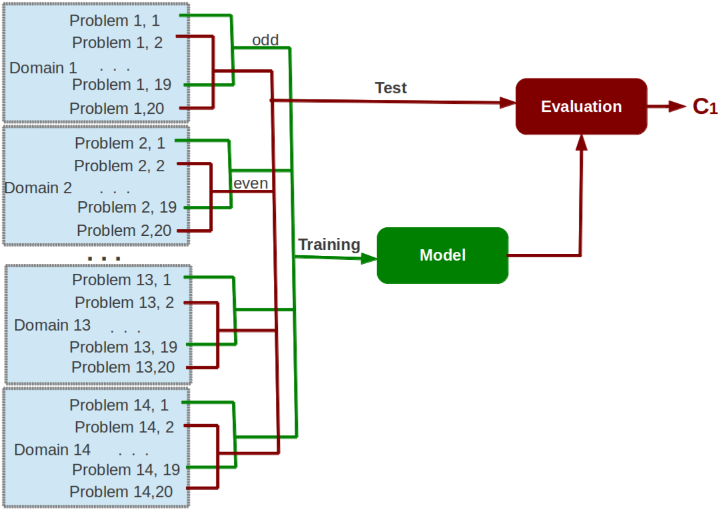
   Regression rules M5Rules
   Instance-based learning algorithms (KNN) [IBk $k = 1, 3, 5$]
   Support vector machine SMOreg

Evaluation

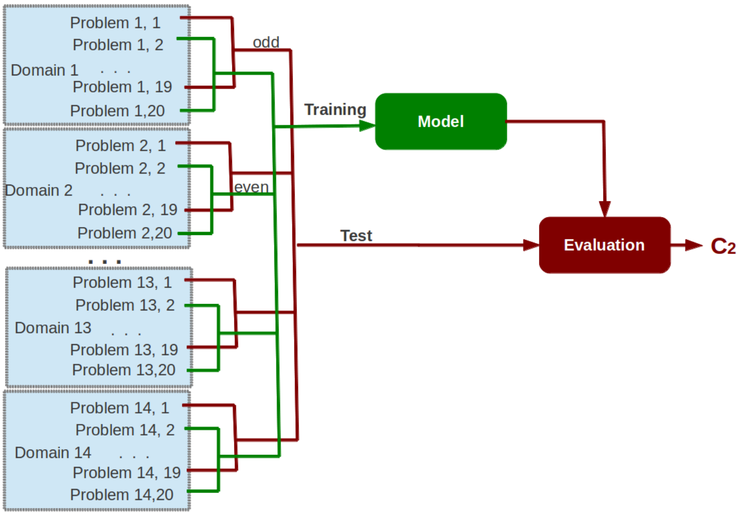
We define two different evaluations:

1. Split evaluation: for problems in known domains
2. Leave one domain out evaluation: for problems in unknown domains

Split Evaluation



Split Evaluation



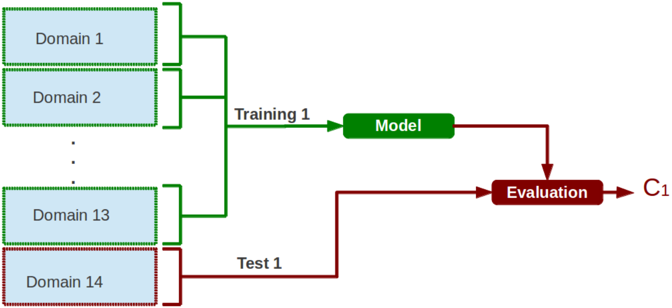
Split Evaluation

Results

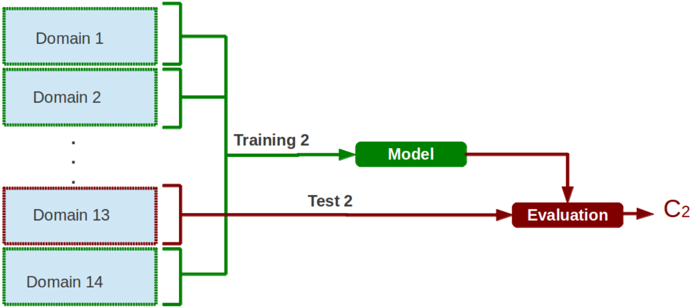
$$R = \frac{1}{2} \sum_{n=1}^{2} C_n$$

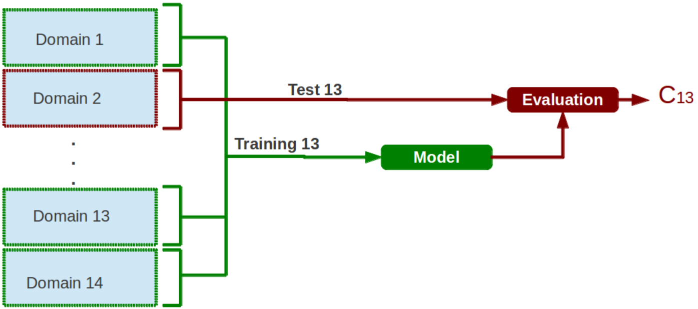
- R: the global results
- C_n: the results in each cycle

Leave One Domain Out Evaluation

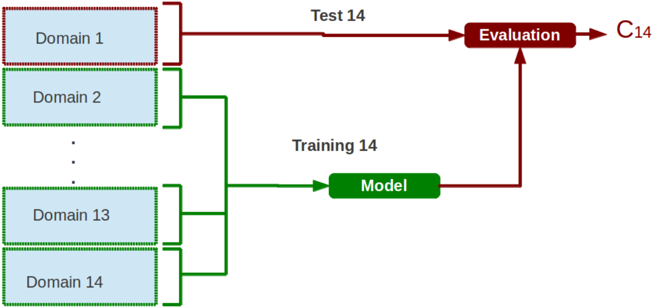


Leave One Domain Out Evaluation




Domain 1
Domain 2
Domain 13
Domain 14
Test 13
Training 13
Model
Evaluation
C13

Leave One Domain Out Evaluation



Leave One Domain Out Evaluation

Results

$$R = \frac{1}{14} \sum_{n=1}^{14} C_n$$

- R: the global results
- C_n: the results in each domain

Classification Results

Given a problem *p* in a domain *d*, will the planner *pl* be able to find a solution in 1800 seconds? (yes-no question → binary classification)

Accuracy results

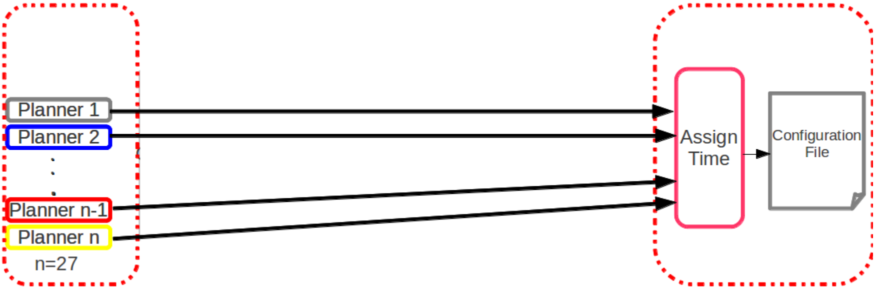
Data set	Split Validation	Leave One Domain Out
J48	**88.75 (1.05)**	59.14 (12.13)
IBk -K 1	88.67 (1.29)	60.83 (10.13)
IBk -K 3	87.63 (1.07)	60.58 (11.76)
IBk -K 5	88.58 (1.07)	**61.95 (11.10)**
SMO	72.48 (1.58)	61.34 (10.10)

Regression Results

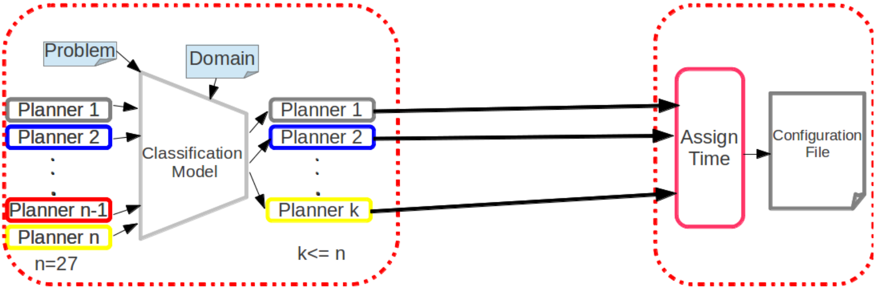
Given a problem *p* in a domain *d*, how long (in seconds) will the planner *pl* take for finding the best solution?
Relative absolute error results

Algorithm	Split validation	Leave one Domain Out
M5Rules	73.66 (3.61)	985.64 (2200.93)
IBk -K 1	67.57 (4.07)	93.66 (23.38)
IBk -K 3	**62.98 (3.12)**	85.96 (22.26)
IBk -K 5	64.39 (3.00)	**85.57 (19.21)**
SMOreg	69.50 (2.87)	907.32 (2620.74)

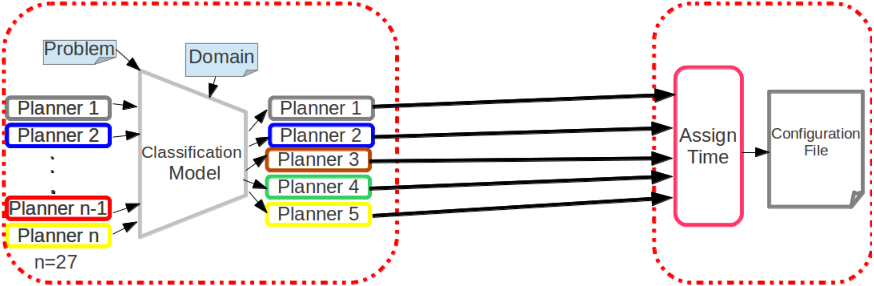
Equal Time


Planner 1
Planner 2
Planner n-1
Planner n
n=27
Assign Time
Configuration File

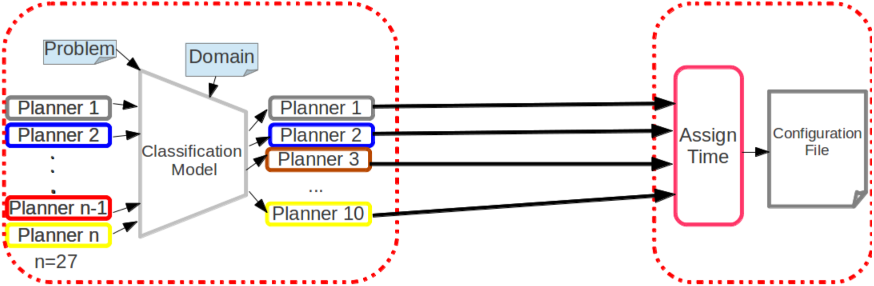
Informed strategy - Best Confidence Estimation(BCE)



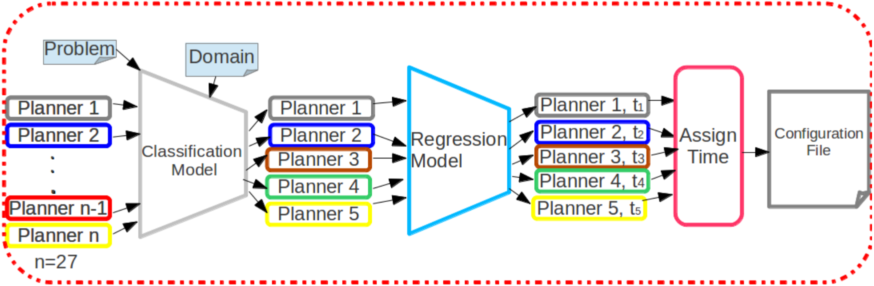
Informed strategy - Best 5 Confidence (B5C)



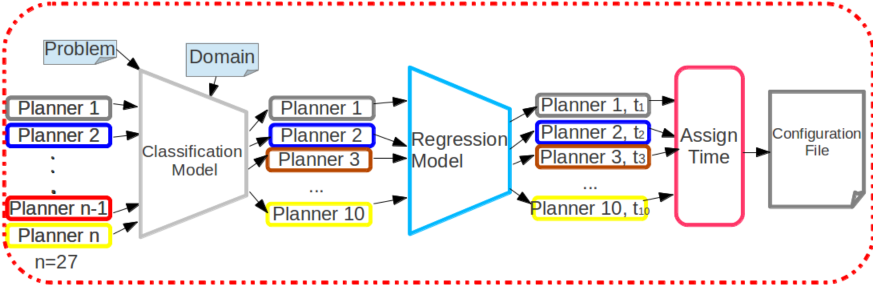
Informed strategy - Best 10 Confidence (B10C)



Informed strategy - Best 5 Regression (B5R)



Informed strategy - Best 10 Regression (B10R)



Experimental set up

- All planners in sequential satisfying track (27 planners)
- All domains in the same track (14 domains)
- Two evaluations (split and leave one domain out evaluation)
- One strategy without knowledge and five strategies with knowledge.
- The results are compared with lama-2011
- There is an upper bound to the best possible strategy (S)

Split Evaluation: Coverage

	ET	BCE	B5C	B10C	B5R	B10R	Lama11	S
Barman	20	20	20	20	20	20	20	20
Elevators	20	20	20	20	20	20	20	20
Floortile	8	8	8	8	8	8	6	**9**
Nomystery	15	18	17	17	18	17	10	**19**
Openstacks	20	20	20	20	20	20	20	20
Parcprinter	20	20	20	20	20	20	20	20
Parking	12	20	20	20	20	20	20	20
Pegsol	20	20	20	20	20	20	20	20
Scanalyzer	18	19	18	17	18	18	**20**	**20**
Sokoban	17	18	**19**	18	**19**	**19**	**19**	**19**
Tibybot	16	18	19	18	17	17	19	**20**
Transport	**20**	**20**	19	**20**	19	**20**	16	**20**
Visitall	20	20	20	20	20	20	20	20
Woodworking	20	20	20	20	20	20	20	20
Total	246	**261**	260	258	259	259	250	**267**

Split Evaluation: Quality improve over Lama

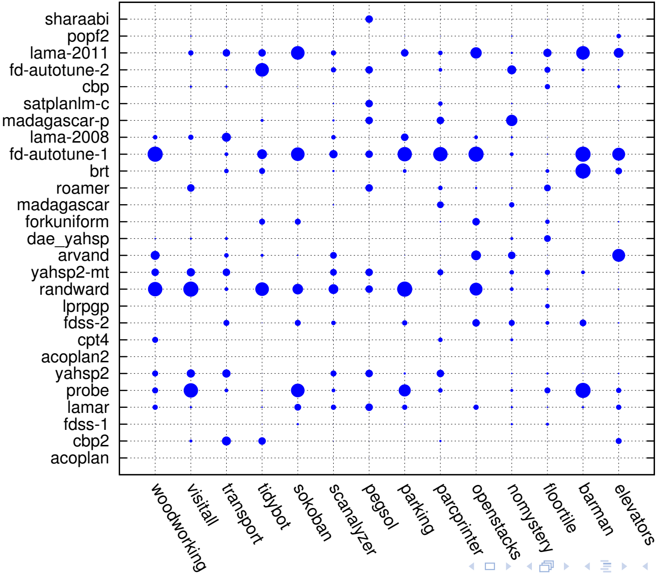
	ET		BCE		B5C		B10C		S
	+	−	+	−	+	−	+	−	+
Barman	**19**	**0**	12	8	18	0	**19**	**0**	**20**
Elevators	16	2	14	6	17	2	**20**	**0**	20
Floortile	4	4	**4**	**0**	4	1	**4**	**0**	5
Nomystery	7	0	**9**	**0**	8	1	8	0	10
Openstacks	2	18	3	6	**5**	**6**	3	9	**17**
Parcprinter	0	20	8	2	8	1	**11**	**0**	11
Parking	3	16	0	20	1	16	4	12	9
Pegsol	0	8	**0**	**2**	**0**	**2**	**0**	**2**	0
Scanalyzer	2	14	9	5	8	4	8	6	**13**
Sokoban	5	10	2	6	4	1	4	2	**6**
Tibybot	5	9	6	5	6	7	**7**	**4**	13
Transport	9	11	11	9	10	8	**14**	**6**	18
Visitall	20	0	18	1	**20**	**0**	**20**	**0**	**20**
Woodworking	9	0	16	2	18	0	**19**	**0**	**19**
Total	101	112	112	72	127	49	**141**	**41**	**181**

Split Evaluation: Quality improve over Lama

	ET		B5R		B10R		S
	+	−	+	−	+	−	+
Barman	**19**	**0**	**19**	**0**	**19**	**0**	**20**
Elevators	16	2	16	2	19	0	20
Floortile	4	4	4	2	**4**	**0**	5
Nomystery	7	0	9	1	8	0	10
Openstacks	2	18	4	7	3	11	**17**
Parcprinter	0	20	8	1	**11**	**0**	11
Parking	3	16	1	16	**4**	**13**	9
Pegsol	0	8	**0**	**2**	**0**	**2**	0
Scanalyzer	2	14	8	6	**10**	**3**	**13**
Sokoban	5	10	4	1	**5**	**1**	**6**
Tibybot	5	9	4	7	6	4	13
Transport	9	11	10	8	**14**	**6**	18
Visitall	**20**	**0**	**20**	**0**	**20**	**0**	**20**
Woodworking	9	0	18	0	**19**	**0**	19
Total	101	112	125	53	**142**	**40**	**181**

Leave One Domain Out Evaluation: Coverage

	ET	BCE	B5C	B10C	B5R	B10R	Lama11	S
Barman	20	20	20	20	20	20	20	20
Elevators	**20**	**20**	17	**20**	18	**20**	**20**	**20**
Floortile	8	**9**	6	**9**	6	**9**	6	**9**
Nomystery	15	17	13	15	13	15	10	**19**
Openstacks	**20**	1	**20**	**20**	15	15	**20**	**20**
Parcprinter	20	20	20	20	20	20	20	20
Parking	12	20	20	20	20	20	20	20
Pegsol	20	20	20	20	20	20	20	20
Scanalyzer	18	17	17	17	18	17	**20**	**20**
Sokoban	17	**19**	18	**19**	18	**19**	**19**	**19**
Tibybot	16	18	19	17	15	16	19	**20**
Transport	**20**	13	16	13	13	13	16	**20**
Visitall	**20**	10	10	**20**	10	**20**	**20**	**20**
Woodworking	20	20	20	20	20	20	20	20
Total	246	224	236	**250**	226	244	250	**267**



sharaabi
popf2
lama-2011
fd-autotune-2
cbp
satplanlm-c
madagascar-p
lama-2008
fd-autotune-1
brt
roamer
madagascar
forkuniform
dae_yahsp
arvand
yahsp2-mt
randward
lprpgp
fdss-2
cpt4
acoplan2
yahsp2
probe
lamar
fdss-1
cbp2
acoplan
woodworking
visitall
transport
tidybot
sokoban
scanalyzer
pegsol
parking
parcprinter
openstacks
nomystery
floortile
barman
elevators

Conclusions

- We demonstrate that predictive models have excellent results in known domains
- The predictive models are quite good to predict problems in previously unseen domains
- We have defined a set of strategies to configure the portfolio
- We have evaluated them with the problems of the IPC-2011 with two strategies

Future work

- Learn better models for unknown domains
  - learning with more domains
  - creating new features that characterize the problems
  - selection of the planners a priori

Learning Predictive Models to Configure Planning Portfolios

Isabel Cenamor
Tomás de la Rosa
Fernando Fernández

June 11, 2013

Universidad Carlos III de Madrid
[plg]

Fawcett, C., Helmert, M., Hoos, H., Karpas, E., Roger, G., and Seipp, J. (2011).
Fd-autotune: Domain-specific configuration using fast downward.
In *Booklet of the 7th International Planning Competition*.

Gerevini, A., Saetti, A., and Vallati, M. (2009).
An automatically configurable portfolio-based planner with macro-actions: PbP.
In *Proceedings of the 19th International Conference on Automated Planning and Scheduling (ICAPS-09)*.

Gerevini, A., Saetti, A., and Vallati, M. (2011).
Pbp2: Automatic configuration of a portfolio-based multi-planner.
The 2011 International Planning Competition.

Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., and Witten, I. H. (2009).
The weka data mining software: an update.
ACM SIGKDD Explorations Newsletter, 11(1):10–18.

Accuracy

$$\frac{(Correctly\ \ Classified\ \ Instances)}{Total\ \ instances} * 100$$

Absolute Relative Error

$$\frac{(Observed\ \ value - Accepted\ \ value)}{Accepted\ \ value}$$

Leave one domain out Evaluation: Quality improve over Lama I

	ET		BCE		B5C		B10C		S
	+	−	+	−	+	−	+	−	+
Barman	**19**	**0**	19	0	19	1	19	0	**20**
Elevators	16	2	16	1	14	4	**18**	**0**	**20**
Floortile	4	4	5	0	0	2	**5**	**0**	5
Nomystery	**7**	**0**	7	2	4	4	4	2	**10**
Openstacks	2	18	1	19	3	17	3	16	**17**
Parcprinter	0	20	**11**	**0**	5	12	**11**	**0**	11
Parking	3	16	4	12	2	12	4	12	**9**
Pegsol	0	8	**0**	**2**	**0**	**2**	**0**	**2**	0
Scanalyzer	2	14	8	4	4	6	4	7	**13**
Sokoban	5	10	5	1	1	7	4	1	6
Tibybot	5	9	5	4	6	3	4	7	**13**
Transport	9	11	12	7	8	7	10	7	**18**
Visitall	20	0	7	13	7	13	20	0	**20**
Woodworking	9	0	**19**	**0**	**19**	**0**	**19**	**0**	19
Total	101	112	119	65	92	90	125	54	**181**

Leave one domain out Evaluation: Quality improve over Lama II

	ET		B5R		B10R		S
	+	−	+	−	+	−	+
Barman	**19**	**0**	18	0	18	0	**20**
Elevators	16	2	15	3	18	1	**20**
Floortile	4	4	0	0	5	1	5
Nomystery	**7**	**0**	4	5	5	1	**10**
Openstacks	2	18	1	19	2	16	**17**
Parcprinter	0	20	5	12	**11**	**0**	11
Parking	3	16	2	13	4	15	**9**
Pegsol	0	8	**0**	**2**	**0**	**2**	0
Scanalyzer	2	14	4	6	4	6	**13**
Sokoban	5	10	2	6	5	1	6
Tibybot	5	9	3	7	3	7	13
Transport	9	11	8	9	8	10	18
Visitall	**20**	**0**	7	13	**20**	**0**	20
Woodworking	9	0	**19**	**0**	**19**	**0**	19
Total	101	112	88	95	122	60	**181**