Seventh International Planning Competition: Deterministic Part

Ángel García-Olaya, Sergio Jiménez, Carlos Linares López

Planning and Learning Group (PLG)
Computer Science Department
Universidad Carlos III de Madrid (Spain)

June 14, 2011

Outline

Disclaimer

From the IPC 2006

Please note that with our evaluation by no means we intend to imply that one particular planner is better than the other ones in general. There can be other evaluation criteria that give different results. Everybody who is interested in the results of IPC 2011 is strongly encouraged to look them and derive her/his own conclusions!

Apology

From the IPC 2008

There are too many planners, domains and results than could be properly introduced in 30 minutes

- IPC website:
  `http://www.plg.inf.uc3m.es/ipc2011-deterministic`
- Domains, planners, summarized data, software and documentation:
  `svn://svn@pleiades.plg.inf.uc3m.es/`*`hidden`*
- All the results (both raw and validated):
  `svn://svn@pleiades.plg.inf.uc3m.es/`*`hidden`*

Acknowledgements

Our most sincere gratitude to:

- To all the participants!
- Bharatranjan Kavuluri
- Frédéric Maris
- Hootan Nakhost
- Ron Petrick
- Nir Lipovetzky
- Amanda and Andrew Coles
- Óscar Pérez, Jaime Pons and Roberto Fuentes
- Héctor L. Palacios
- Bhaskara Marthi
- Guy Shani
- Tomás de la Rosa
- Daniel Borrajo
- Derek Long
- Daniel L. Kovacs
- To the IPC council and Malte Helmert in particular

Outline

Competition Setting

One of the main drivers has been providing *continuity* with the previous competition:

- All tracks clearly defined before the competition
- The scoring rules clearly stated before the competition
- Blind planner submission
- Explicit requirement to publish source code
- . . . and papers —to be published with the consent of authors
- Strong emphasis on plan quality but measurements have been taken on a wide variety of parameters
- No language extensions at all!

Planner submission

Also, as in the IPC 2008:

- Planners submitted before knowing the domains
- All experiments run by the organizers
- Domains have been kept secret

Still, bug fixes were possible:

- Participants only saw a sample of the instances where we observed bugs
- We checked all the modifications with `KDiff3`

New in the IPC 2011:

- Planners were allowed to make preliminary experiments over the same premises used to run the competition

The PDDL language

- All competitions until this one introduced changes and revisions to PDDL
- This year, no extensions have been introduced
- Daniel L. Kovacs has submitted two unpublished manuscripts with a specification of the latest revision PDDL 3.1

Competition tracks

In previous competitions there were different versions of the same domains
This year:

- Four competition tracks defined in the call for participation
- Each track is a completely separate competition
- No aggregation of results across tracks
- Just a single domain definition provided to all planners

The IPC 2011 tracks

Four different tracks were announced, but only three took place:

- Sequential:
  - STRIPS + action costs (+ fluents)
  - Objective: minimize action cost (sum of action costs)
- Multi-core:
  - STRIPS + action costs
  - Objective: minimize action cost (sum of action costs)
  - The planner is alowed to use all cores simultaneously
- Temporal:
  - STRIPS + action costs (+ numerical fluents)
  - Objective: minimize total time (makespan)

A satisficing and an optimal subtrack were arranged for the Sequential track

Planners

Planners registered and submitted:

- Sequential satisficing: 29 registered, 27 submitted
- Sequential optimal: 22 registered, 14 submitted, 2 witdhrawals
- Sequential Multi-core satisficing: 10 registered, 8 submitted
- Temporal satisficing: 12 registered, 8 submitted
- Temporal optimal: 3 registered, cancelled
- Preferences satisficing: 4 registered, 1 submitted, cancelled
- Preferences optimization: 6 registered, 1 submitted, cancelled

55 planners in total from 55 different people from 11 countries: Australia, Canada, China, France, Germany, India, Israel, Italy, Spain, UK and USA

Ranking the planners

Compliant with the ranking schema of the IPC 2008:

- 12–14 domains per track, 20 tasks each
- Each planner gets a score 0.00–1.00 for each solved task in every domain
- Highest aggregate score wins
- Score only depends on plan quality, but other measurements are available
- Bounds per track: 30 minutes, 6 Gb RAM, 750 Gb HD

Scoring function

- Planners get a score S as a function of the best-known (Q^*) solution and the cost of each generated plan (Q):

$$S = \frac{Q^*}{Q}$$

- No domain-specific solvers have been used
- In the sequential optimization track, all planners produced plans with the same quality

Facilities

As in the IPC 2008, the IPC was organized around three facilities:

- A wiki page to exchange information at different levels among participants and organizers
  http://www.plg.inf.uc3m.es/ipc2011-deterministic
- A svn repository to store domains, planners, software and results
  svn://svn@pleiades.plg.inf.uc3m.es/ipc2011/data
  svn://svn@pleiades.plg.inf.uc3m.es/ipc2011/results
- A cluster of 11 nodes for running the experiments

IPC Software (I)

Most of the tasks have been fully automated:

- IPCData: for building planners, domains, automating the execution and validating results
  - No need to download planners and/or domains
  - However, dumps of our repositories are available for private experiments
- IPCReport: for inspecting data and creating overall score tables
  - $\sim$ 50 different variables can be inspected
  - . . . including memory, runtime, quality plan, # solutions, diagnosis, . . .
- IPCPrivate: code specific to the IPC 2011
  - To enable repeatibility

IPC Software (& II)

All the code used at the IPC 2011 is publicly available

- The source code can be found at:
  `svn://svn@pleiades.plg.inf.uc3m.es/ipc2011/data/scripts`
- On-line documentation is available at:
  `http://www.plg.inf.uc3m.es/sw-ipc2011/`
- Licensed under the GNU Public License version 3

Please use it!

Outline

Sequential Satisficing track (I)

- ACOPlan and ACOPlan2: Marco Baioletti, Alfredo Milani, Valentina Poggioni, Fabio Rossi
- Arvand: Hootan Nakhost, Martin Mueller, Richard Valenzano, Fan Xie
- BRT: Vidal Alcázar, Manuela Veloso
- CBP and CBP2: Raquel Fuentetaja
- Roamer: Qiang Lu, You Xu, Ruoyun Huang, Yixin Chen
- CPT4: Vincent Vidal
- DAE-YAHSP: Johann Dréo, Marc Schoenauer, Pierre Saveant, Vincent Vidal

Sequential Satisficing track (II)

- Fast Downward Autotune-1 and Fast Downward Autotune-2: Chris Fawcett, Malte Helmert, Holger Hoos, Erez Karpas, Gabi Röger, Jendrik Seipp
- Fast Downward Stone Soup-1 and Fast Downward Stone Soup-2: Malte Helmert, Erez Karpas, Silvia Richter, Gabi Röger, Jendrik Seipp
- Fork Uniform: Michael Katz, Carmel Domshlak
- LAMA-2008: Silvia Richter, Matthias Westphal
- LAMA-2011: Silvia Richter, Matthias Westphal, Malte Helmert, Gabi Röger
- Lamar: Alan Olsen, Daniel Bryce
- LPRPG-P: Amanda Coles, Andrew Coles

Sequential Satisficing track (& III)

- Madagascar and Madagascar-p: Jussi Rintanen
- POPF2: Amanda Coles, Andrew Coles, Maria Fox, Derek Long
- Probe: Nir Lipovetzky, Héctor Geffner
- Randward: Alan Olsen, Daniel Bryce
- SATPLANLM-C: Dunbo Cai, Minghao Yin
- Sharaabi: Bharat Ranjan Kavuluri
- YAHSP2 and YAHSP2-MT: Vincent Vidal

Sequential Optimization (I)

- BJOLP: Erez Karpas
- CPT4: Vincent Vidal
- Fast Downward Autotune: Chris Fawcett, Malte Helmert, Holger Hoos, Erez Karpas, Gabi Röger, Jendrik Seipp
- Fast Downward Stone Soup-1 and Fast Downward Stone Soup-2: Malte Helmert, Jörg Hoffmann, Erez Karpas, Emil Keyder, Raz Nissim, Silvia Richter, Gabi Röger, Jendrik Seipp, Matthias Westphal
- Fork Init: Michael Katz, Carmel Domshlak

Sequential Optimal (& II)

- Gamer: Peter Kissmann, Stefan Edelkamp
- IFork Init: Michael Katz, Carmel Domshlak
- LM-cut: Malte Helmert, Carmel Domshlak
- LMFork: Michael Katz, Carmel Domshlak
- Merge and Shrink: Raz Nissim, Malte Helmert, Jörg Hoffmann
- Selective Max: Erez Karpas, Carmel Domshlak, Malte Helmert, Shaul Markovitch

Sequential Satisficing Multi-core

- ACOPlan: Marco Baioletti, Alfredo Milani, Valentina Poggioni, Fabio Rossi
- Arvand Herd: Hootan Nakhost, Martin Mueller, Jonathan Schaeffer, Nathan Sturtevant, Richard Valenzano
- ay-Also-Plan Threaded: Juhan Ernits, Charles Gretton
- Roamer-p: You Xu, Qiang Lu, Ruoyun Huang, Yixin Chen
- Madagascar and Madagascar-p: Jussi Rintanen
- PHSFF: Moisés Martínez
- YAHSP2-MT: Vincent Vidal

Temporal Satisficing

- CPT4: Vincent Vidal
- DAEYAHSP: Johann Dréo, Marc Schoenauer, Pierre Savéant, Vincent Vidal
- LMTD: Yanmei Hu, Dunbo Cai, Minghao Yin
- POPF2: Amanda Coles, Andrew Coles, Maria Fox, Derek Long
- Sharaabi: Bharat Ranjan Kavuluri
- TLP-GP: Frederic Maris, Pierre Regnier
- YAHSP2: Vincent Vidal
- YAHSP2-MT: Vincent Vidal

Outline

Overview

- All domains are STRIPS
- But fluents were used in the temporal track as well and some sequential domains
- New and old domains from the IPC 2008:
  - Sequential tracks: Five new domains which pose new challenges. 14 domains
  - Temporal track: Four new domains with 3 that explicitly require concurrency. 12 domains
- No baseline planner has been used

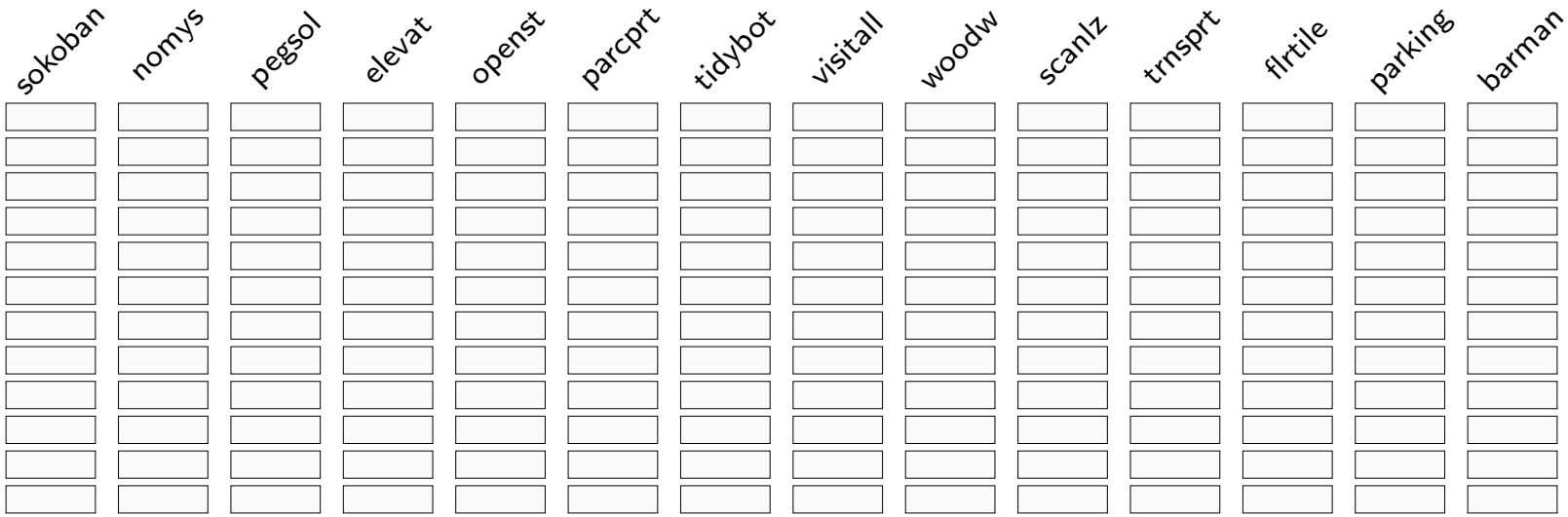
19 different domains in total!

Domains

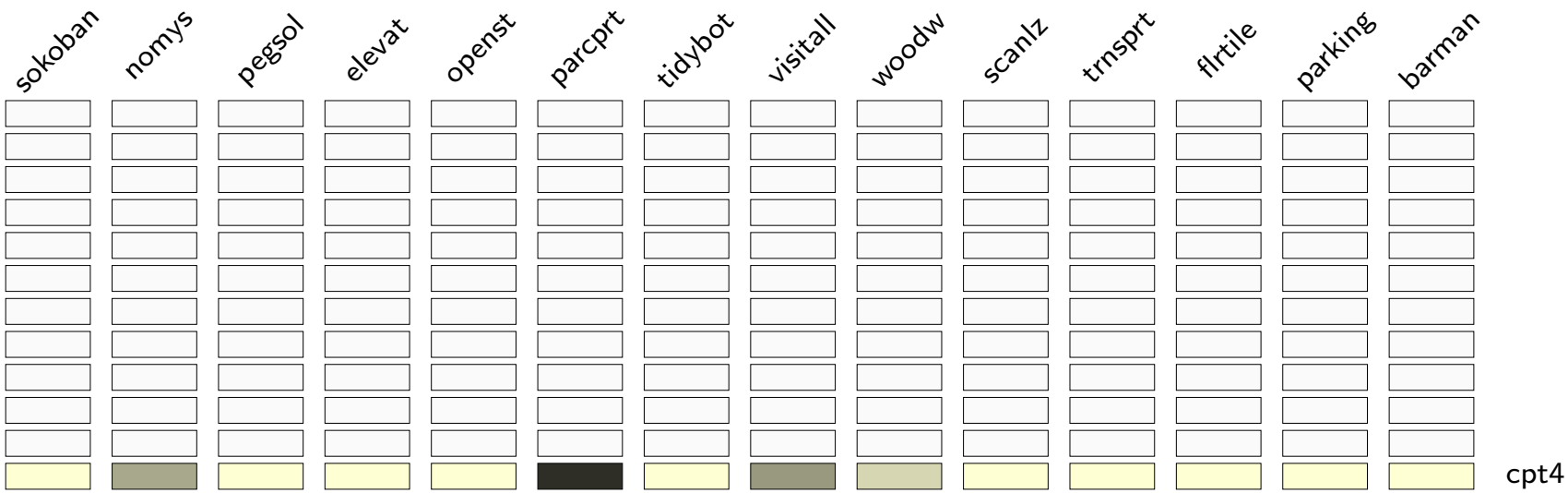
barman	sequential	scanalyzer	sequential
crewplanning	temporal	sokoban	sequential, temporal
elevators	sequential, temporal	storage	temporal
floortile	sequential, temporal	tidybot	sequential
matchcellar	temporal	tms	temporal
nomystery	sequential	turnandopen	temporal
openstacks	sequential, temporal	transport	sequential
parcprinter	sequential, temporal	visitall	sequential
parking	sequential, temporal	woodworking	sequential
pegsol	sequential, temporal		

(*Reused domains in blue and new ones in red*)

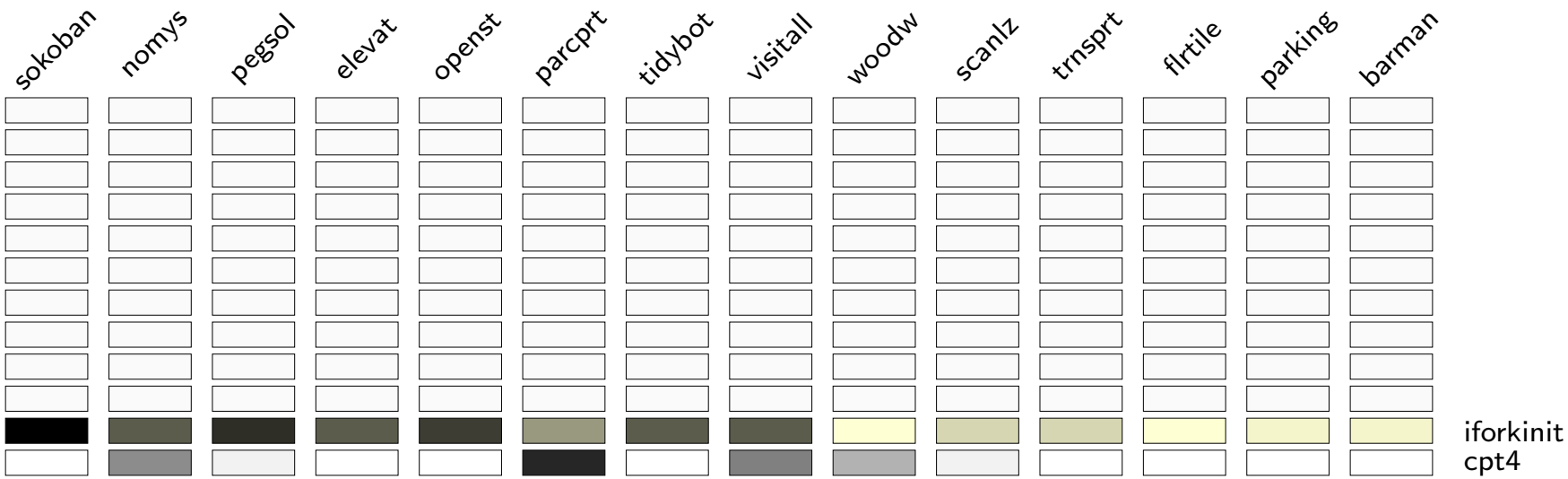
Sequential Optimization track: Results



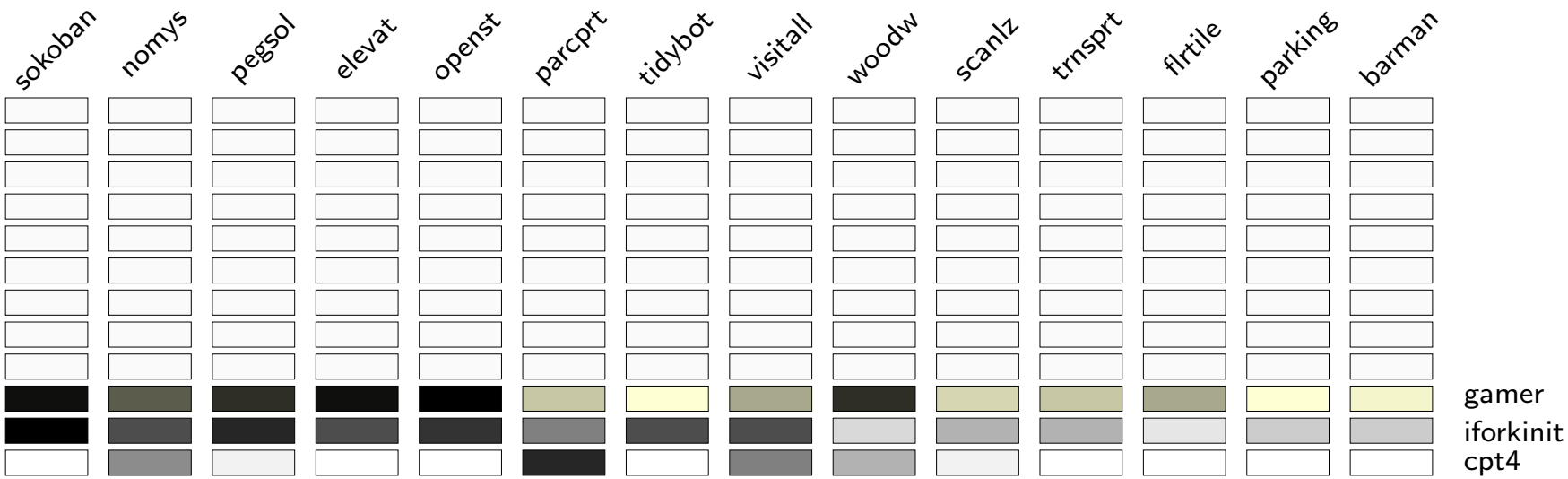
Sequential Optimization track: Results



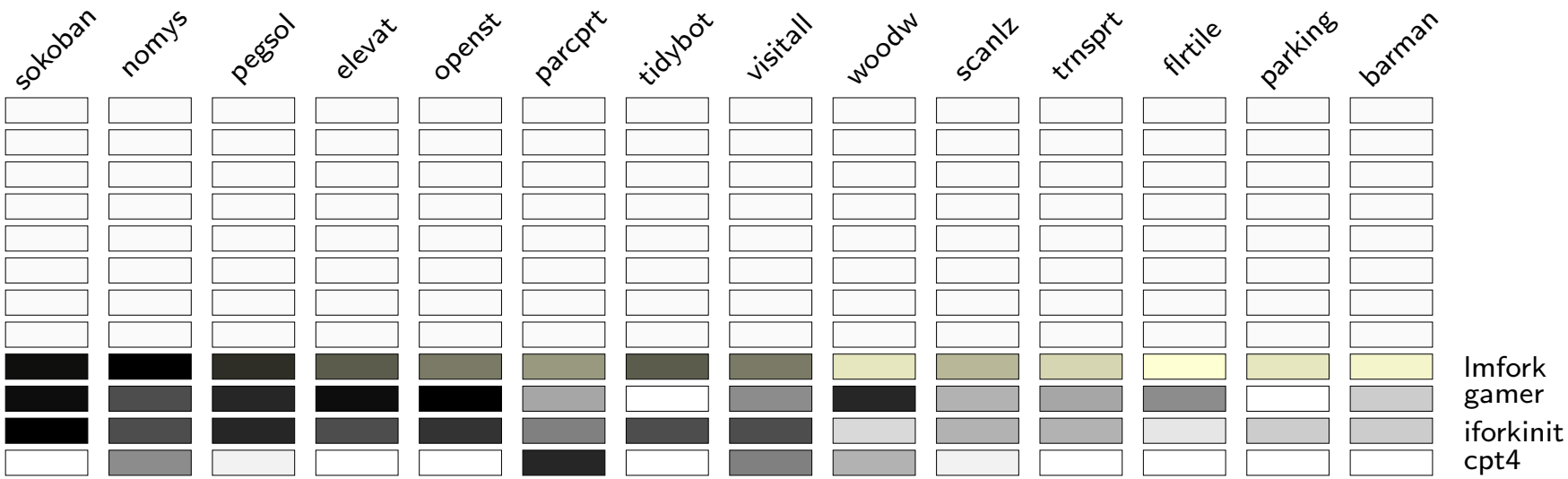
Sequential Optimization track: Results



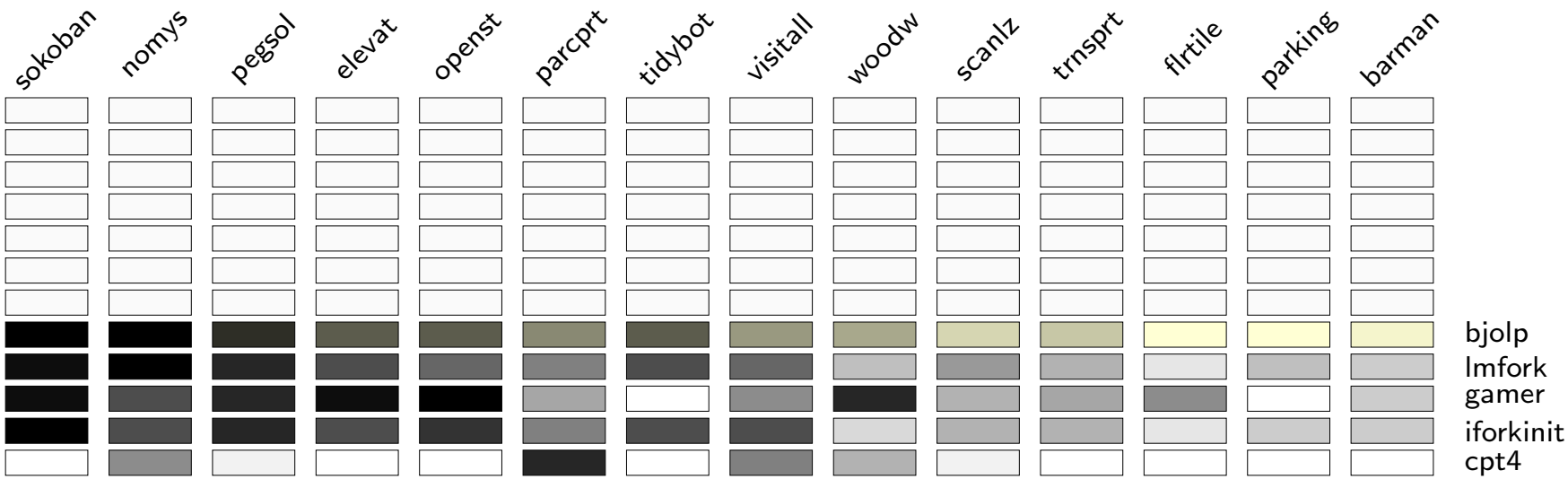
Sequential Optimization track: Results



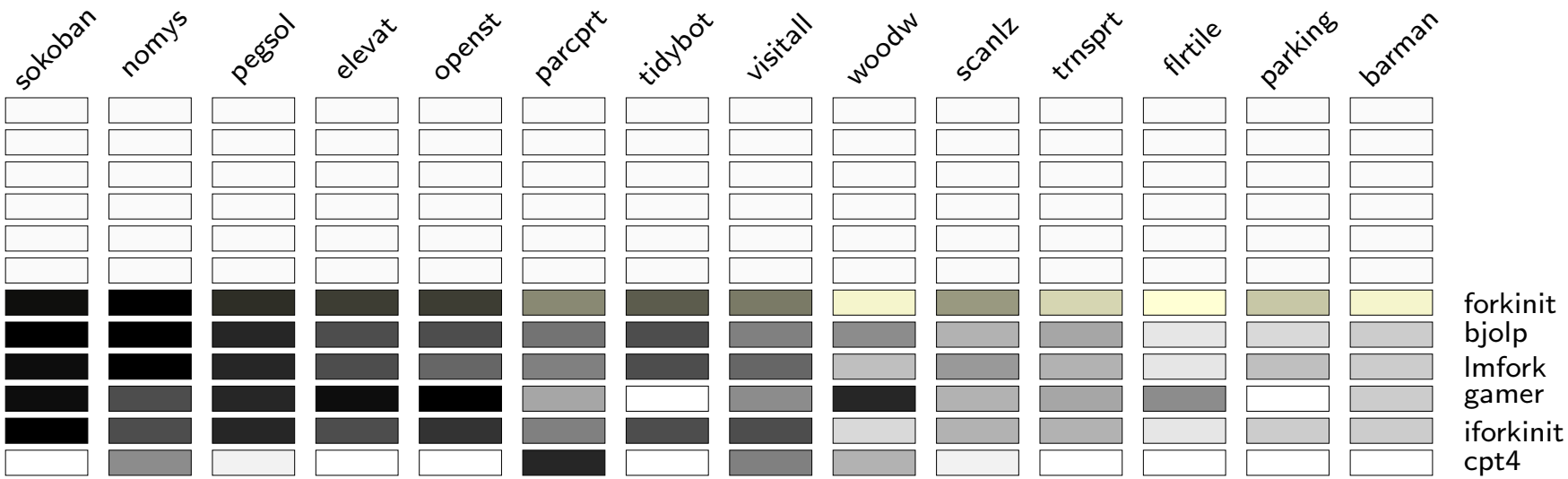
Sequential Optimization track: Results



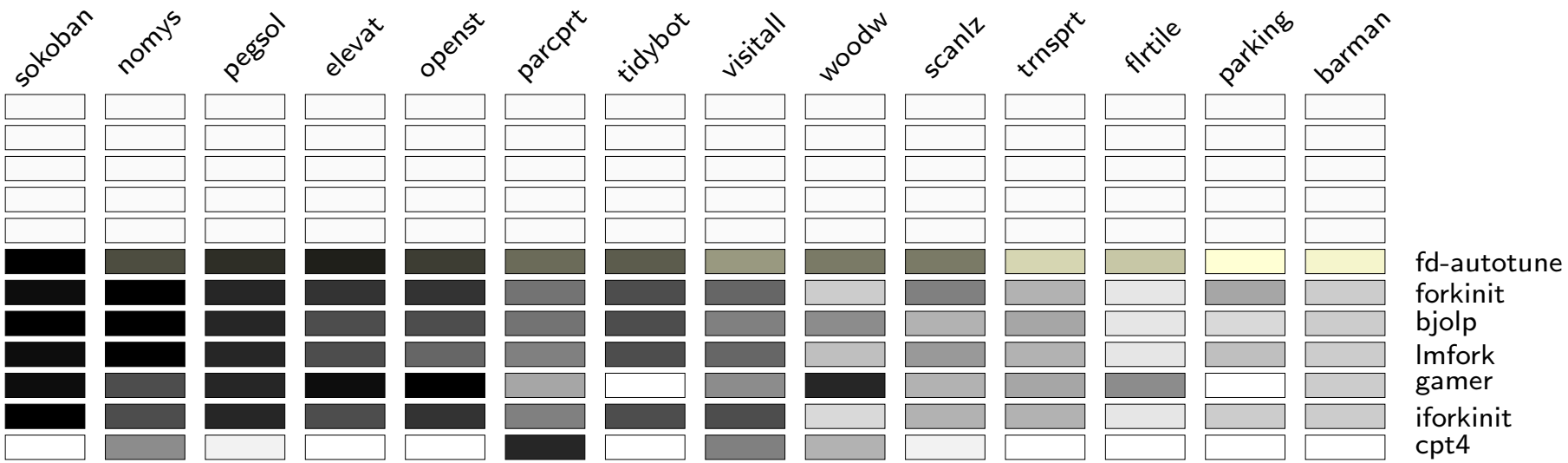
Sequential Optimization track: Results



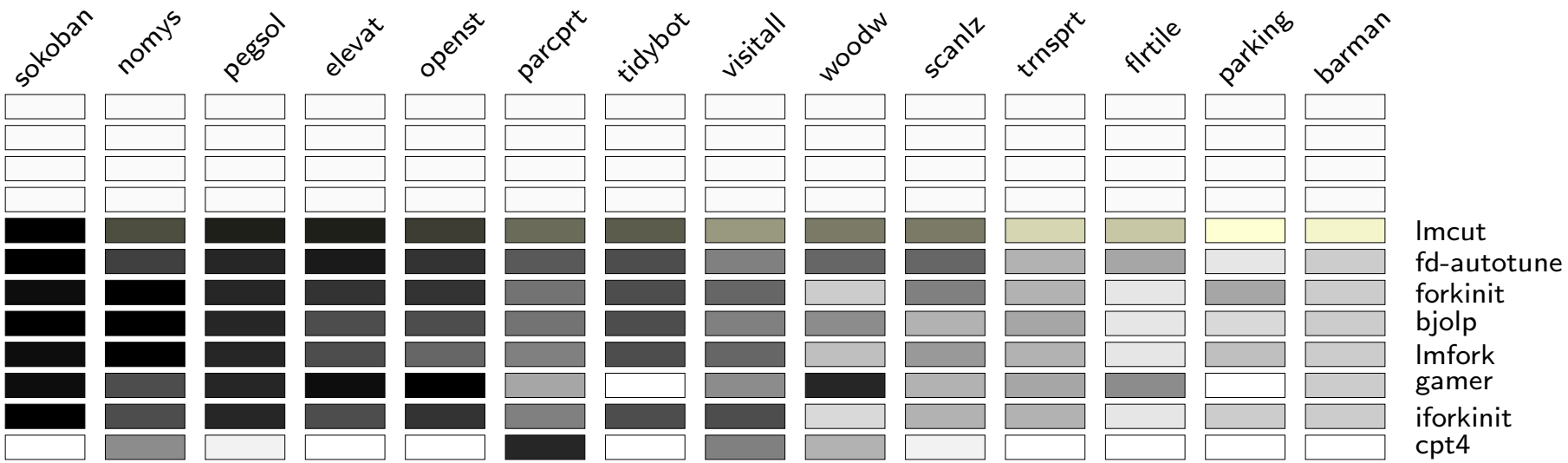
Sequential Optimization track: Results



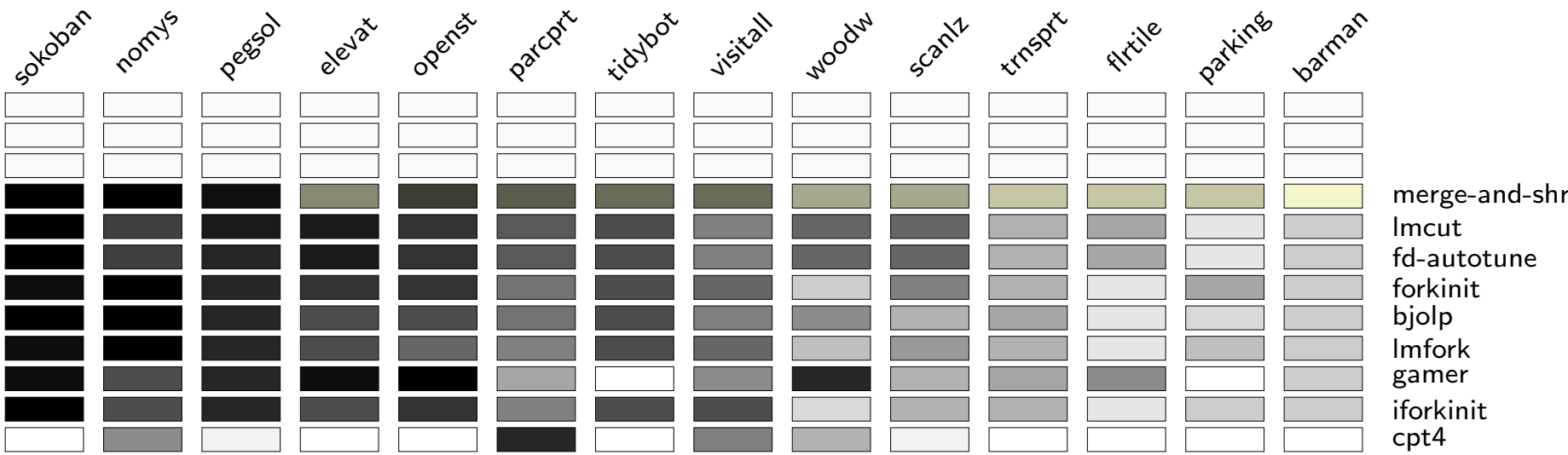
Sequential Optimization track: Results



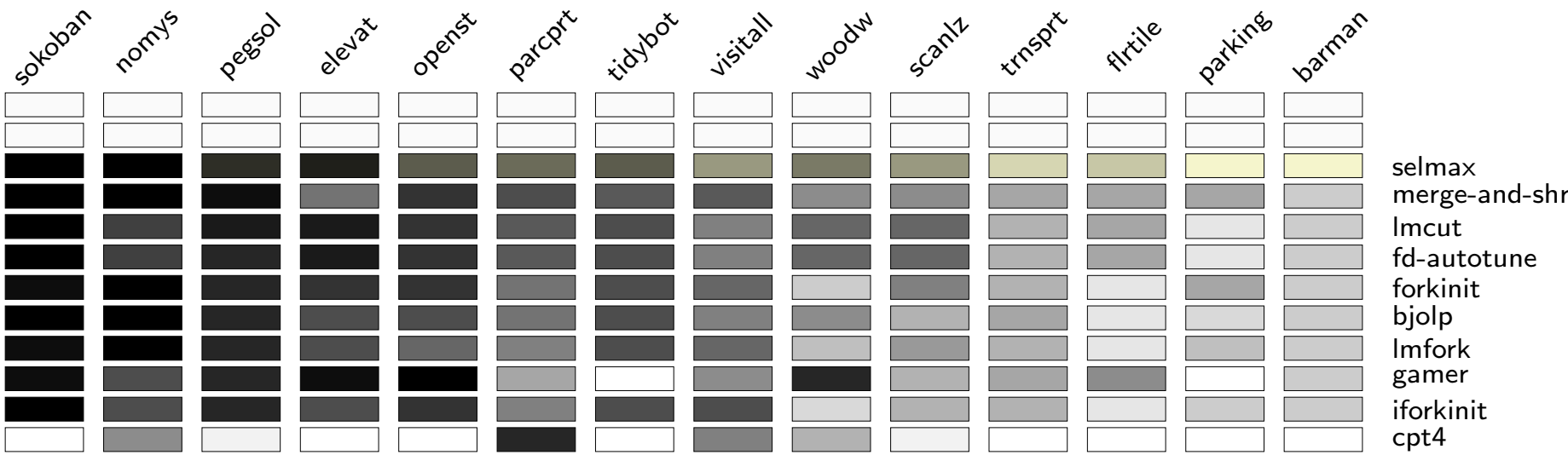
Sequential Optimization track: Results



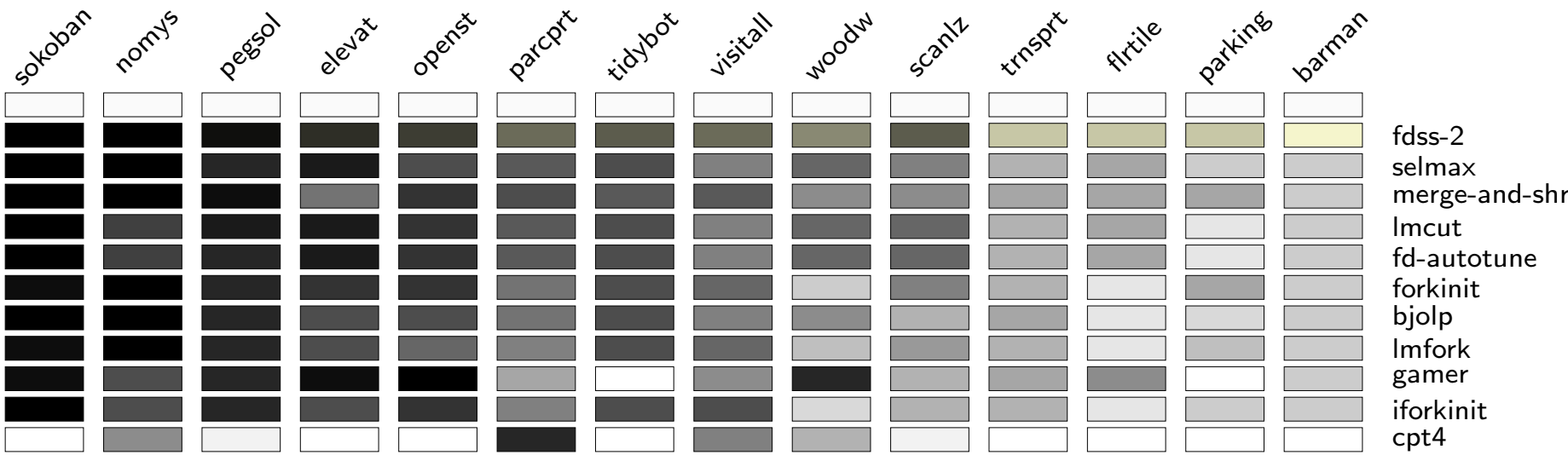
Sequential Optimization track: Results



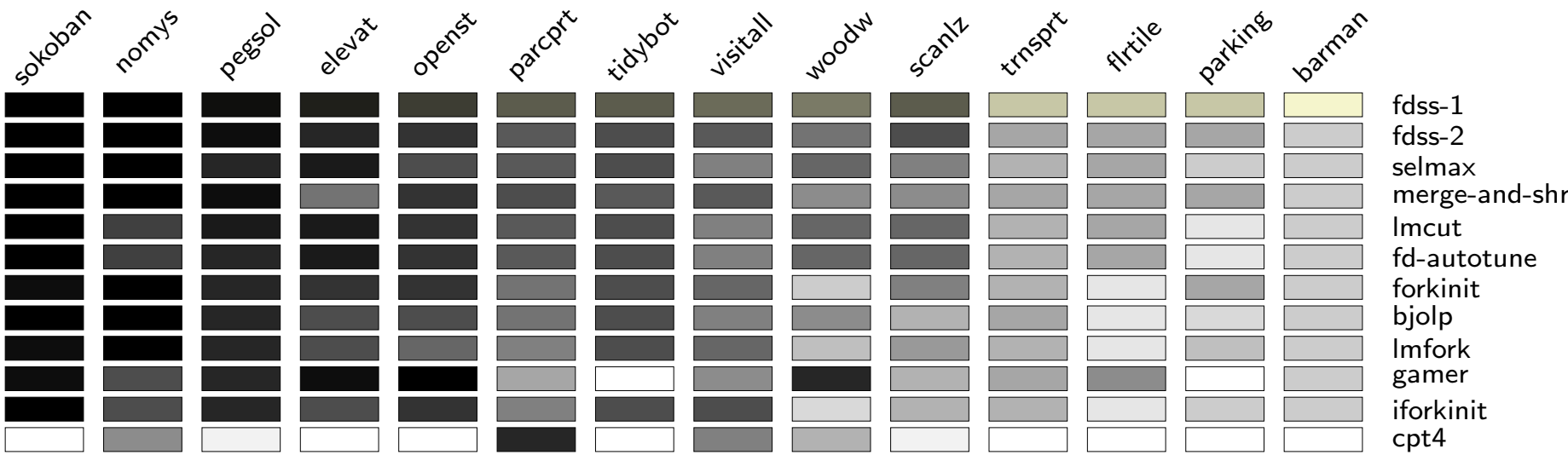
Sequential Optimization track: Results



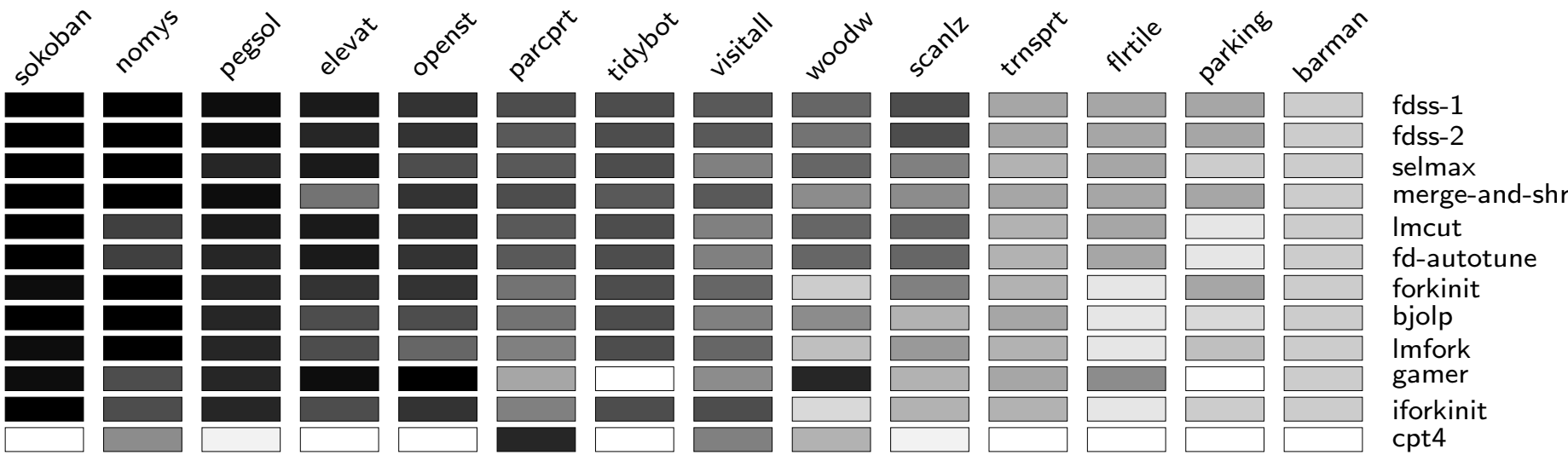
Sequential Optimization track: Results



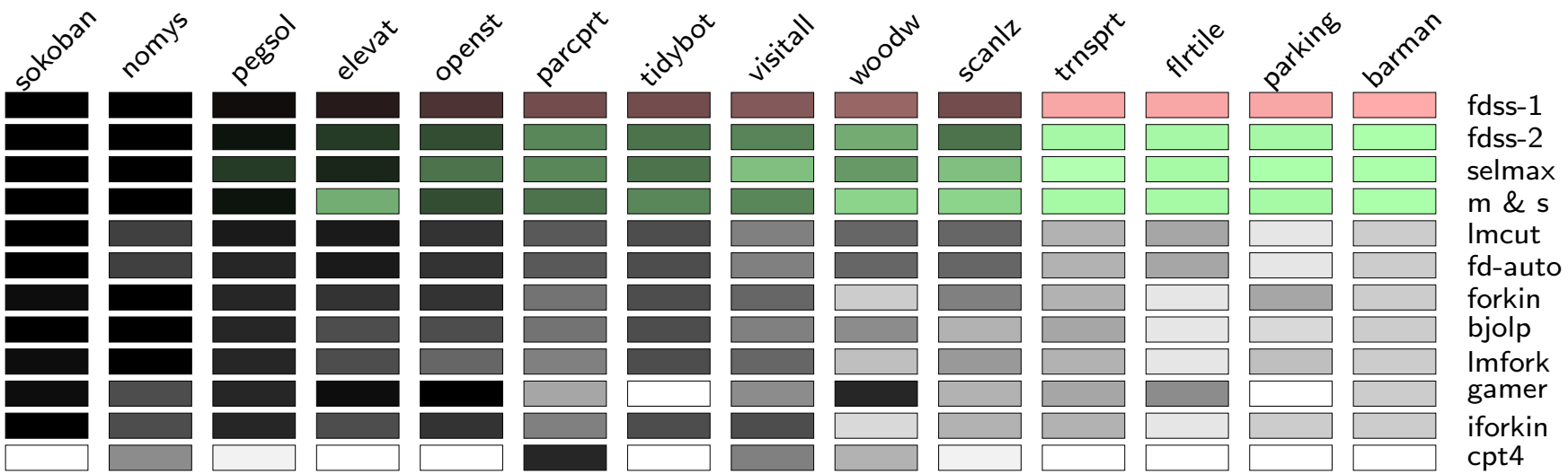
Sequential Optimization track: Results



Sequential Optimization track: Results



Sequential Optimization track: Results



Sequential Optimization track: Awards

Winner

Fast Downward Stone Soup-1: Malte Helmert, Jörg Hoffmann, Erez Karpas, Emil Keyder, Raz Nissim, Silvia Richter, Gabi Röger, Jendrik Seipp, Matthias Westphal

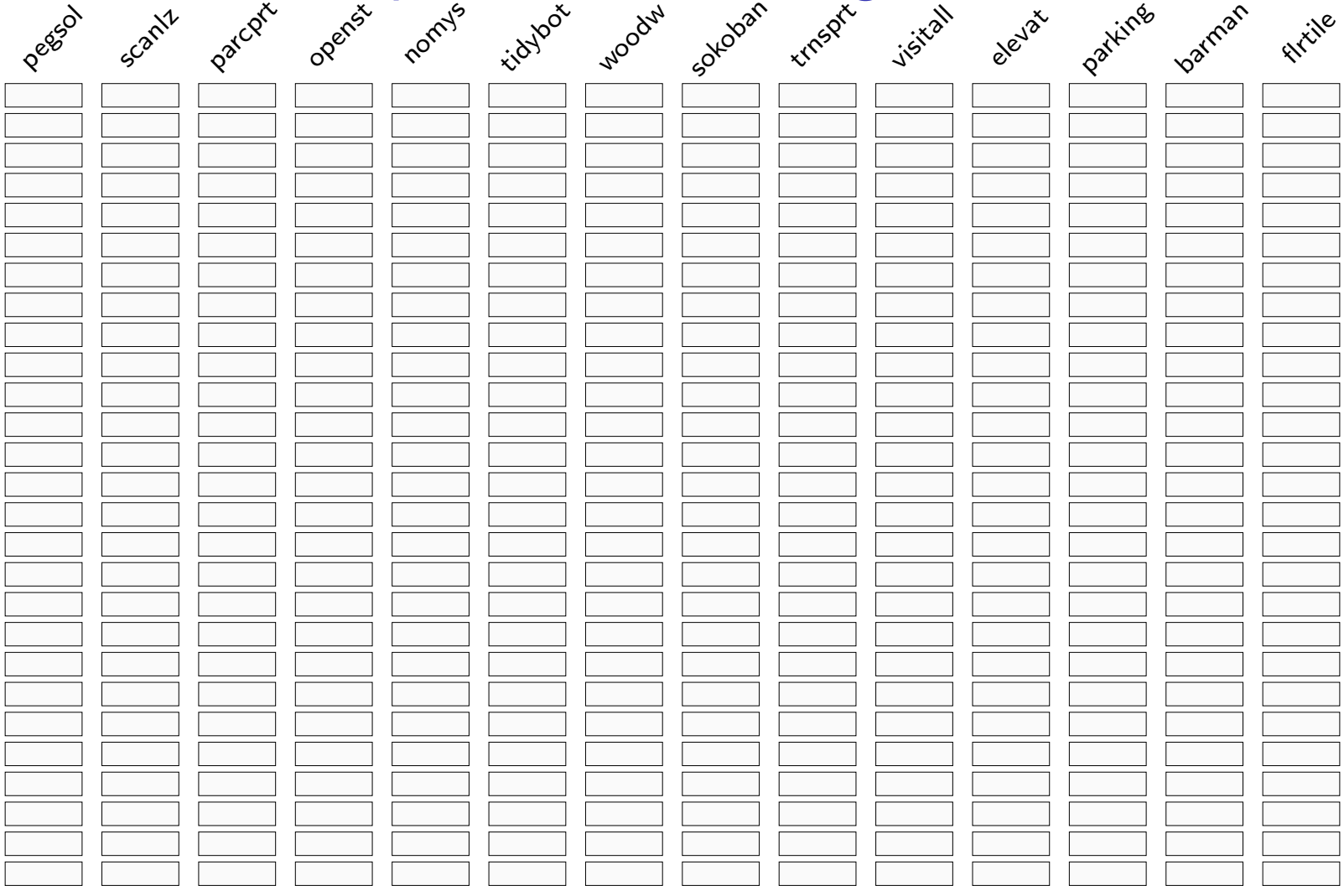
Runner-up ex-aequo

Selective Max: Erez Karpas, Carmel Domshlak, Malte Helmert, Shaul Markovitch

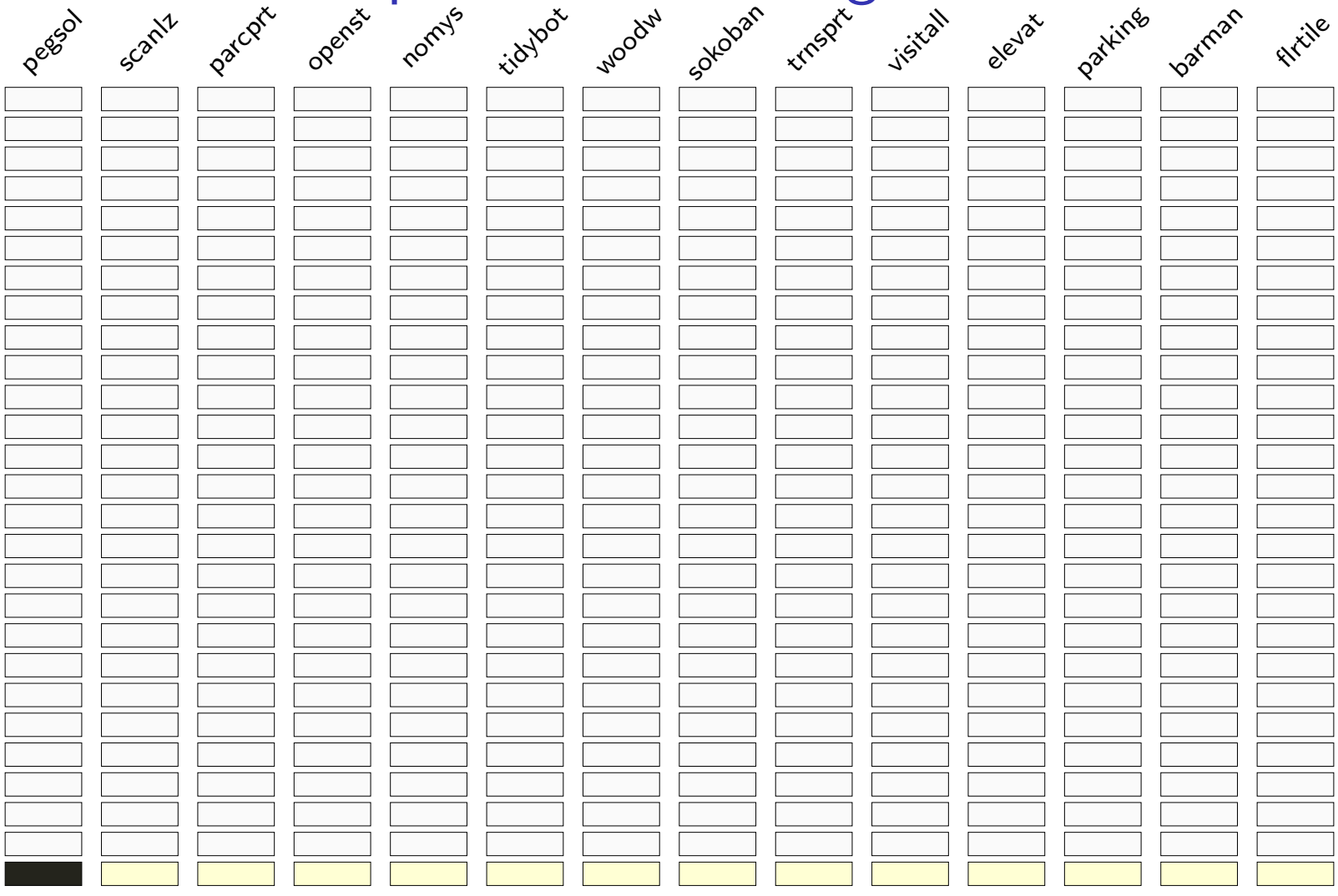
Runner-up ex-aequo

Merge and Shrink: Raz Nissim, Malte Helmert, Jörg Hoffmann

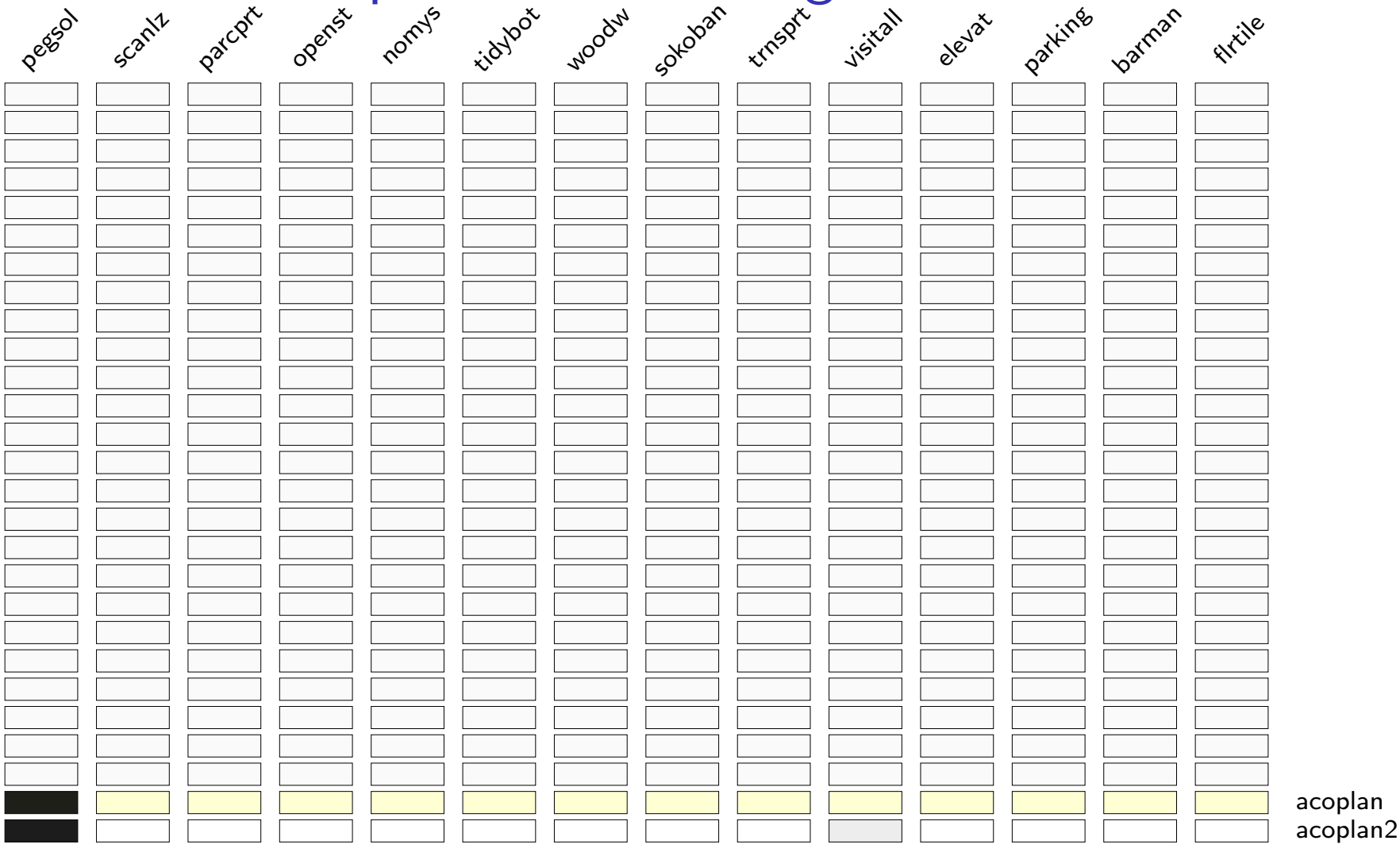
Sequential Satisficing track: Results



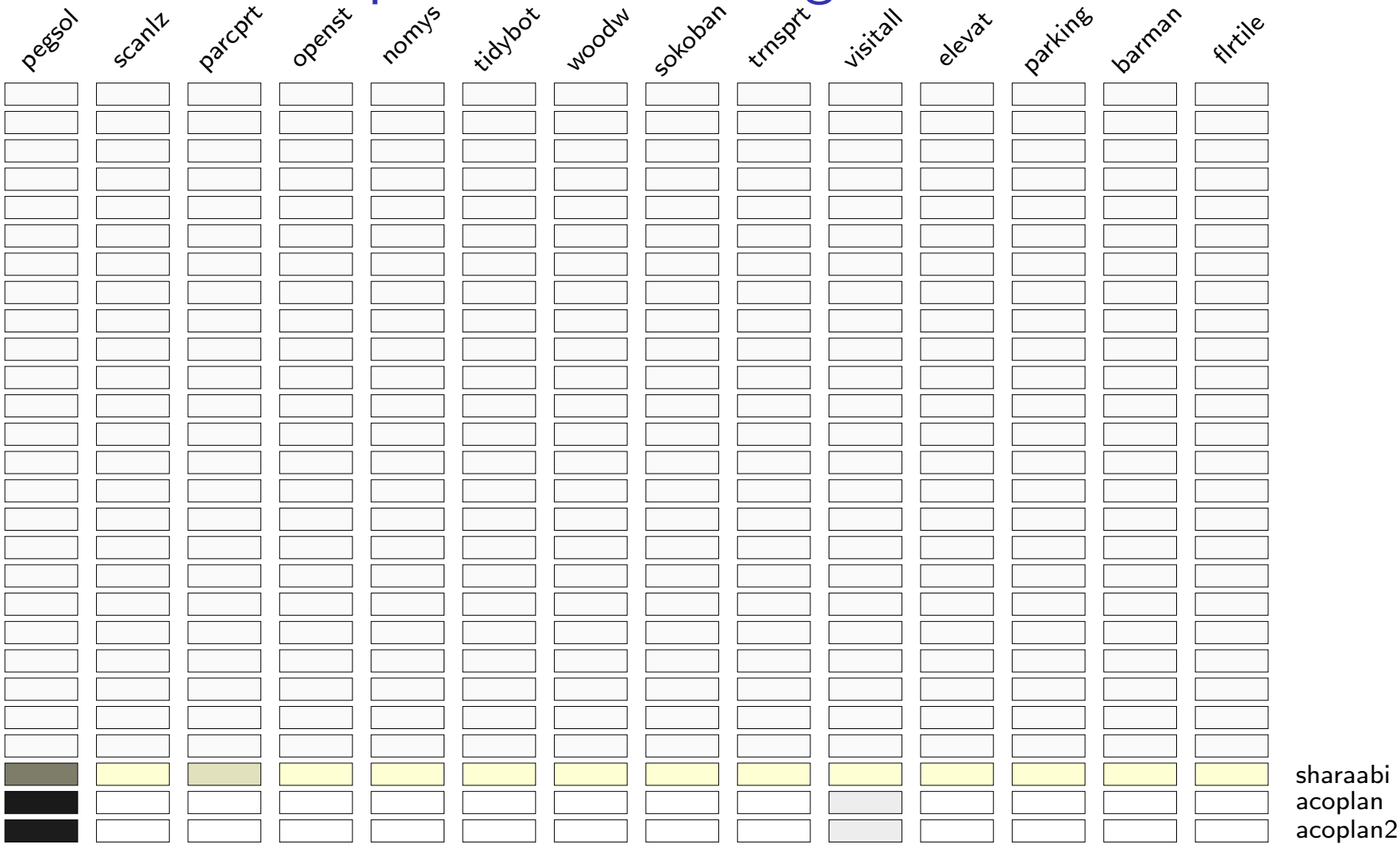
Sequential Satisficing track: Results



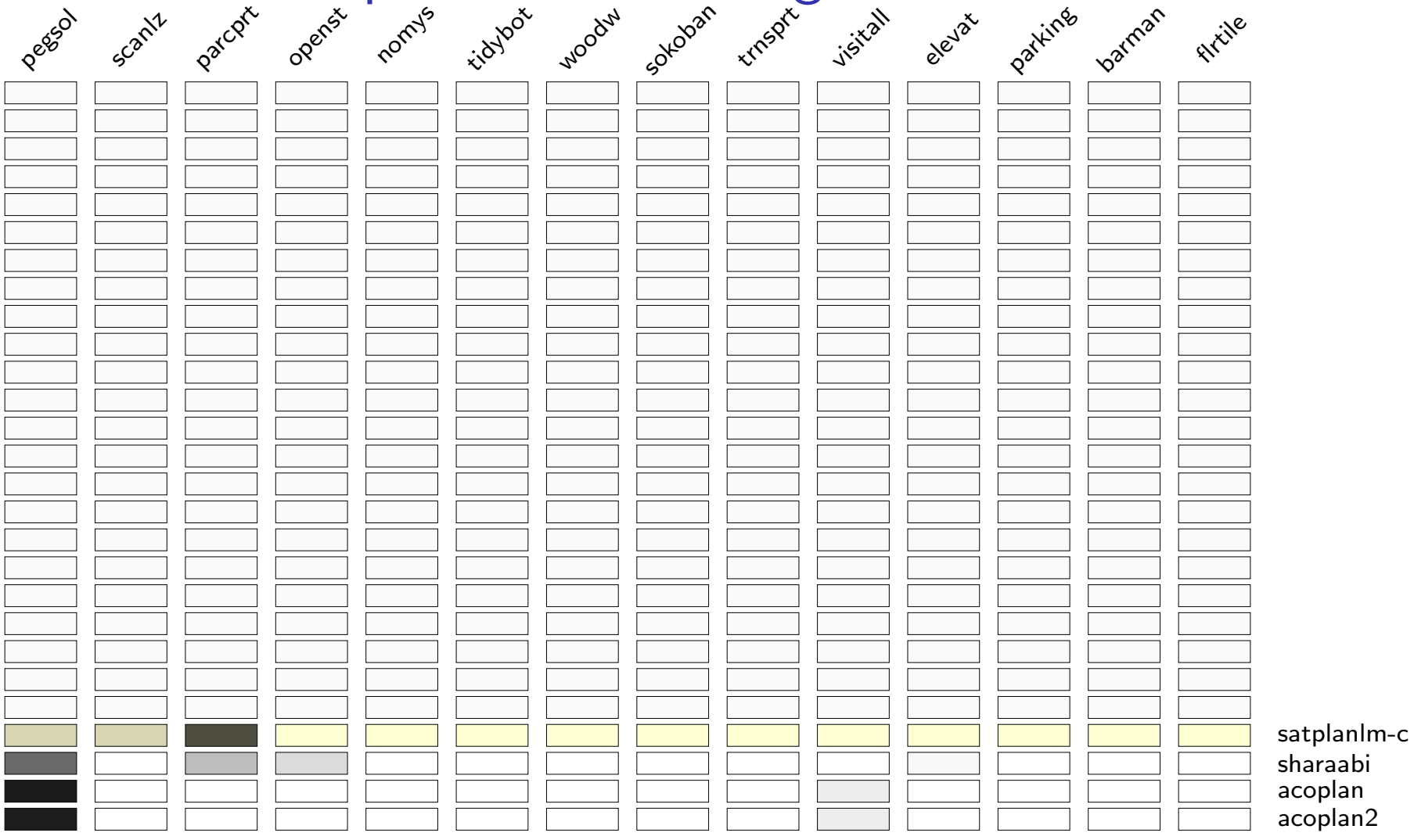
Sequential Satisficing track: Results



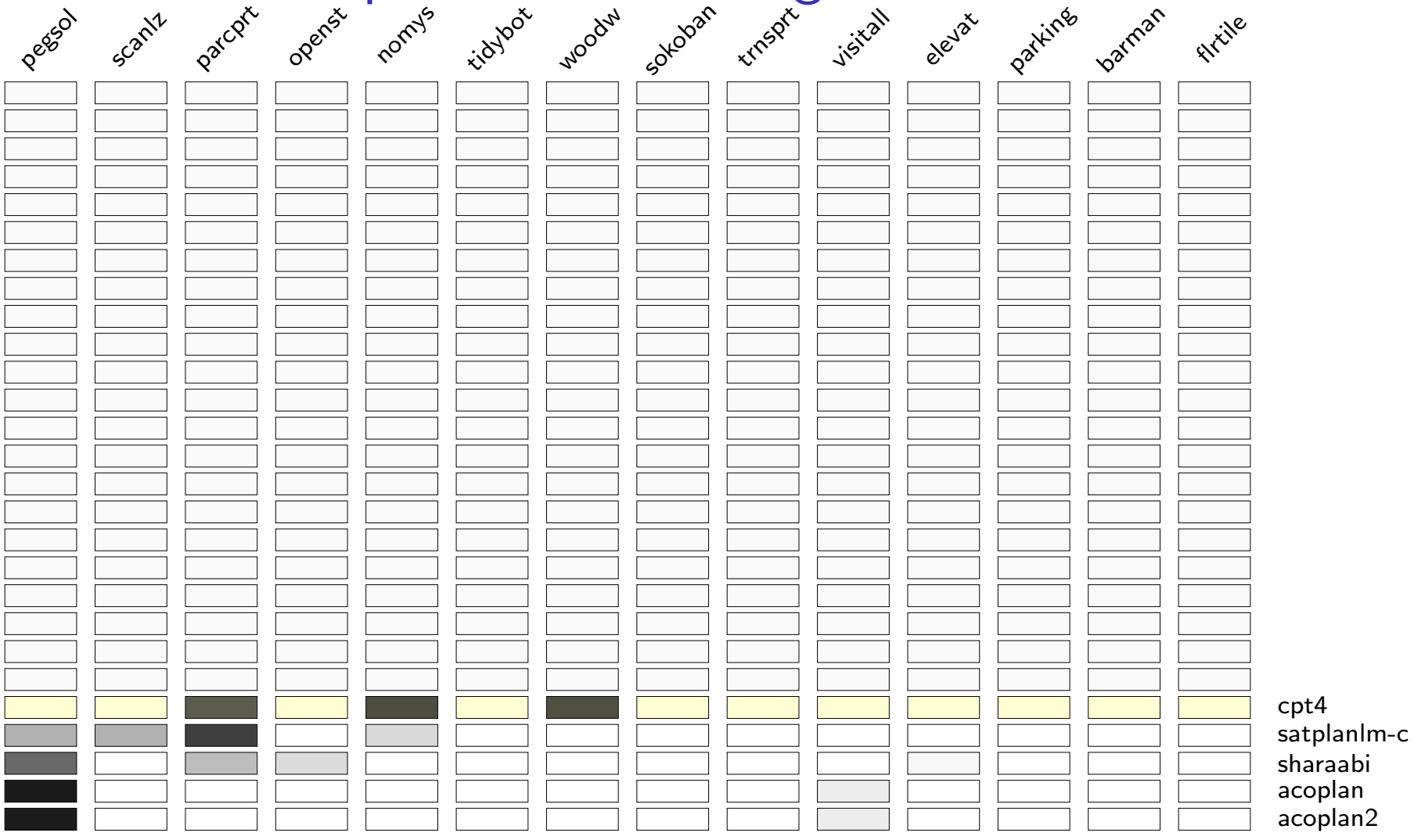
Sequential Satisficing track: Results



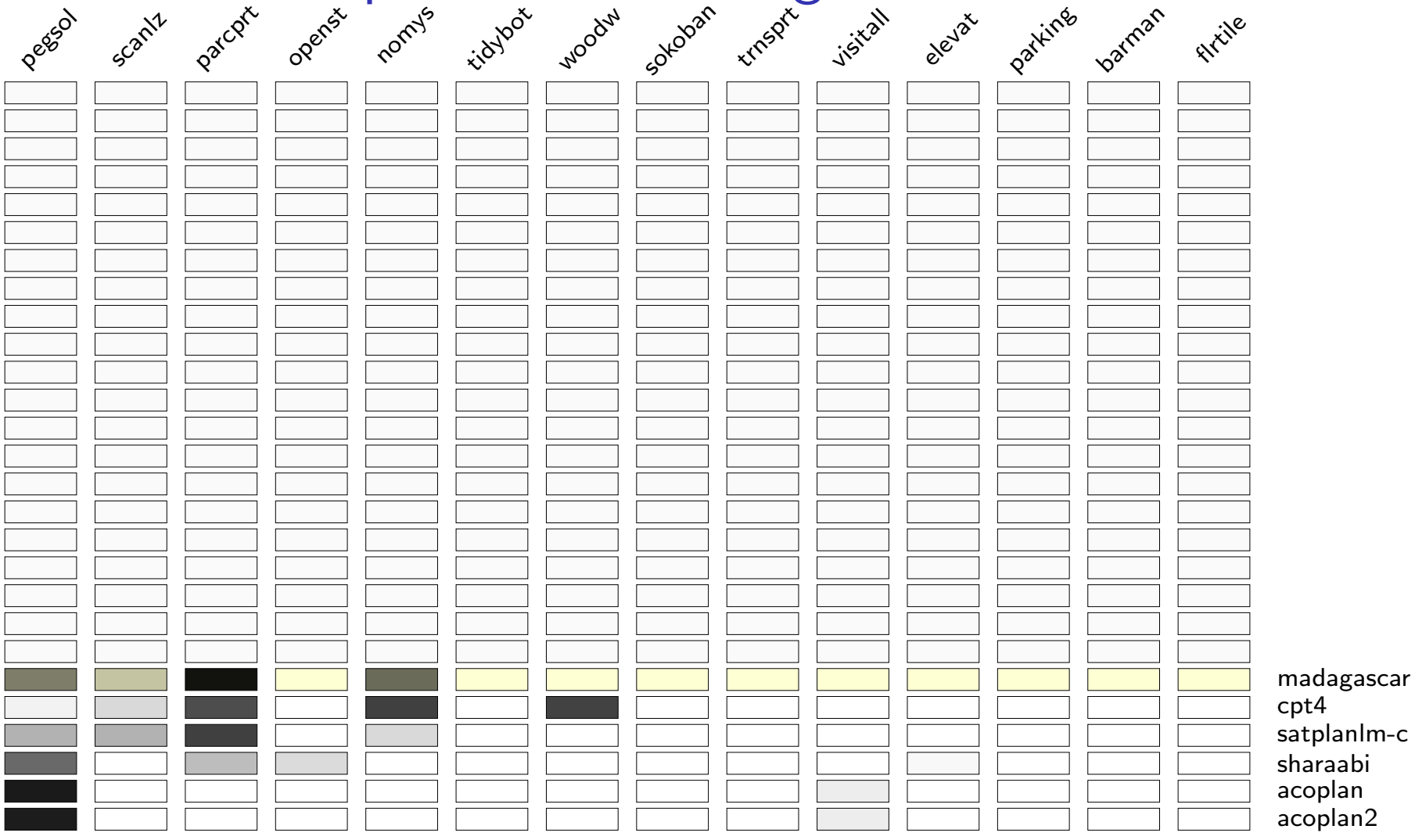
Sequential Satisficing track: Results



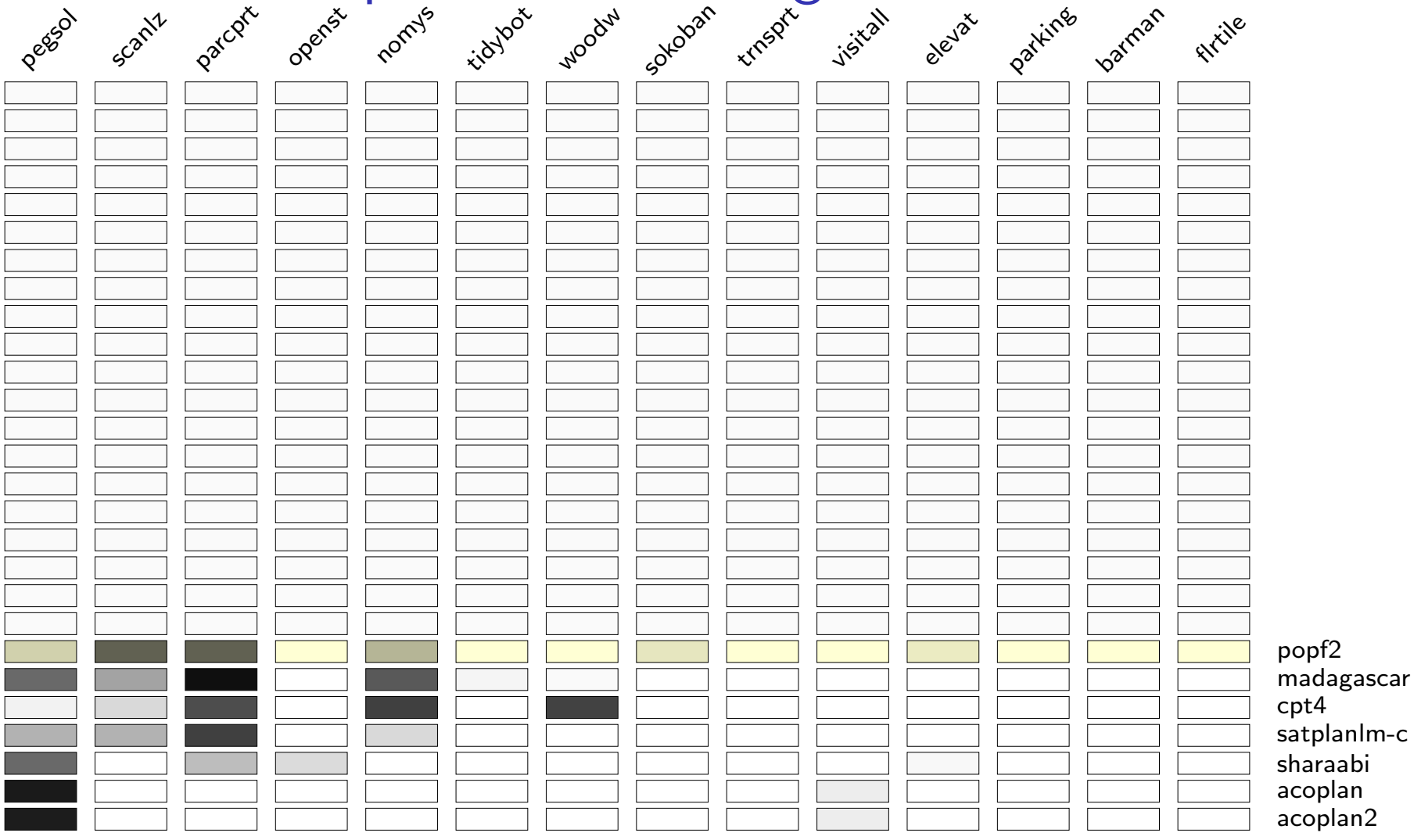
Sequential Satisficing track: Results



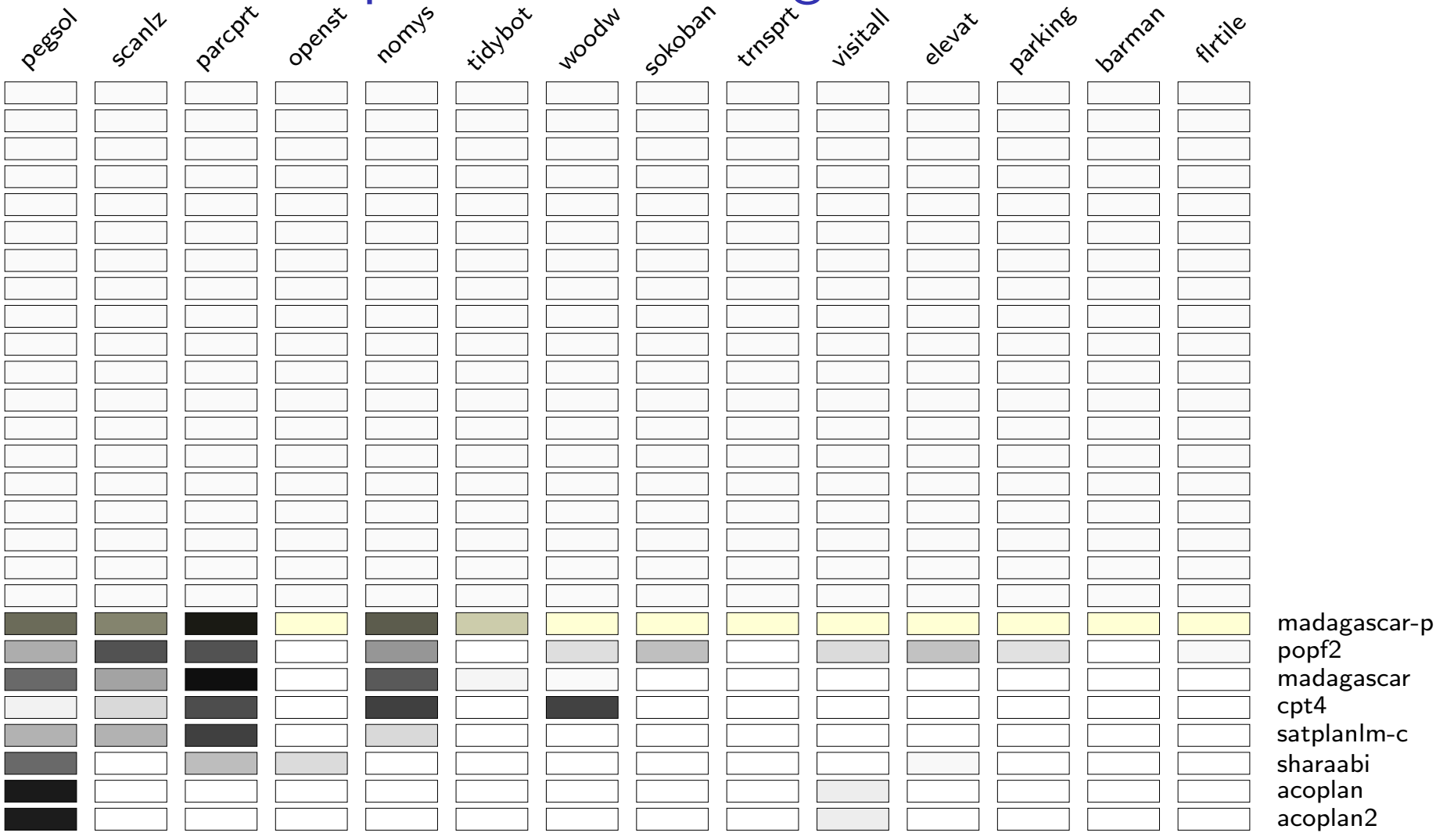
Sequential Satisficing track: Results



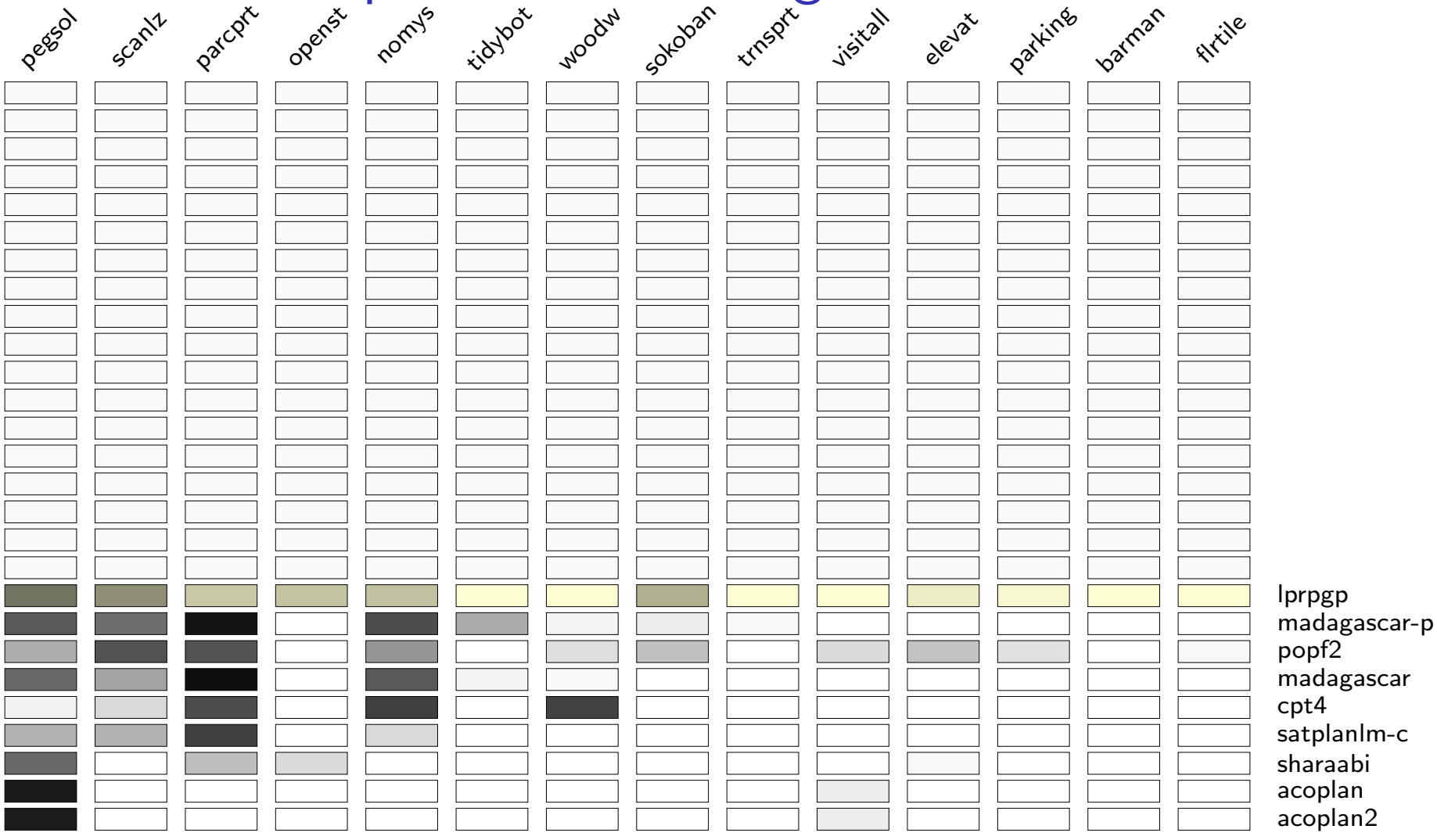
Sequential Satisficing track: Results



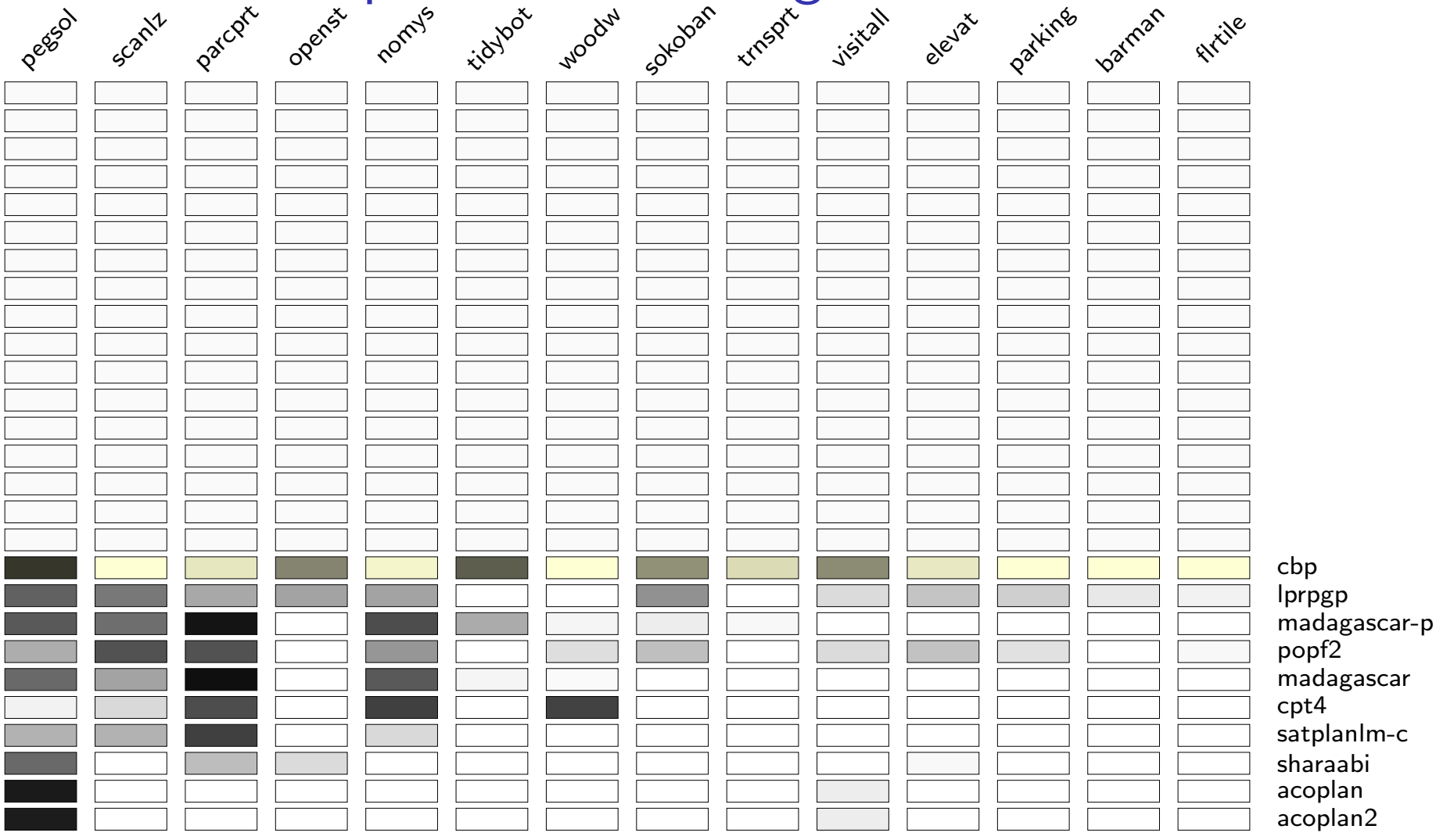
Sequential Satisficing track: Results



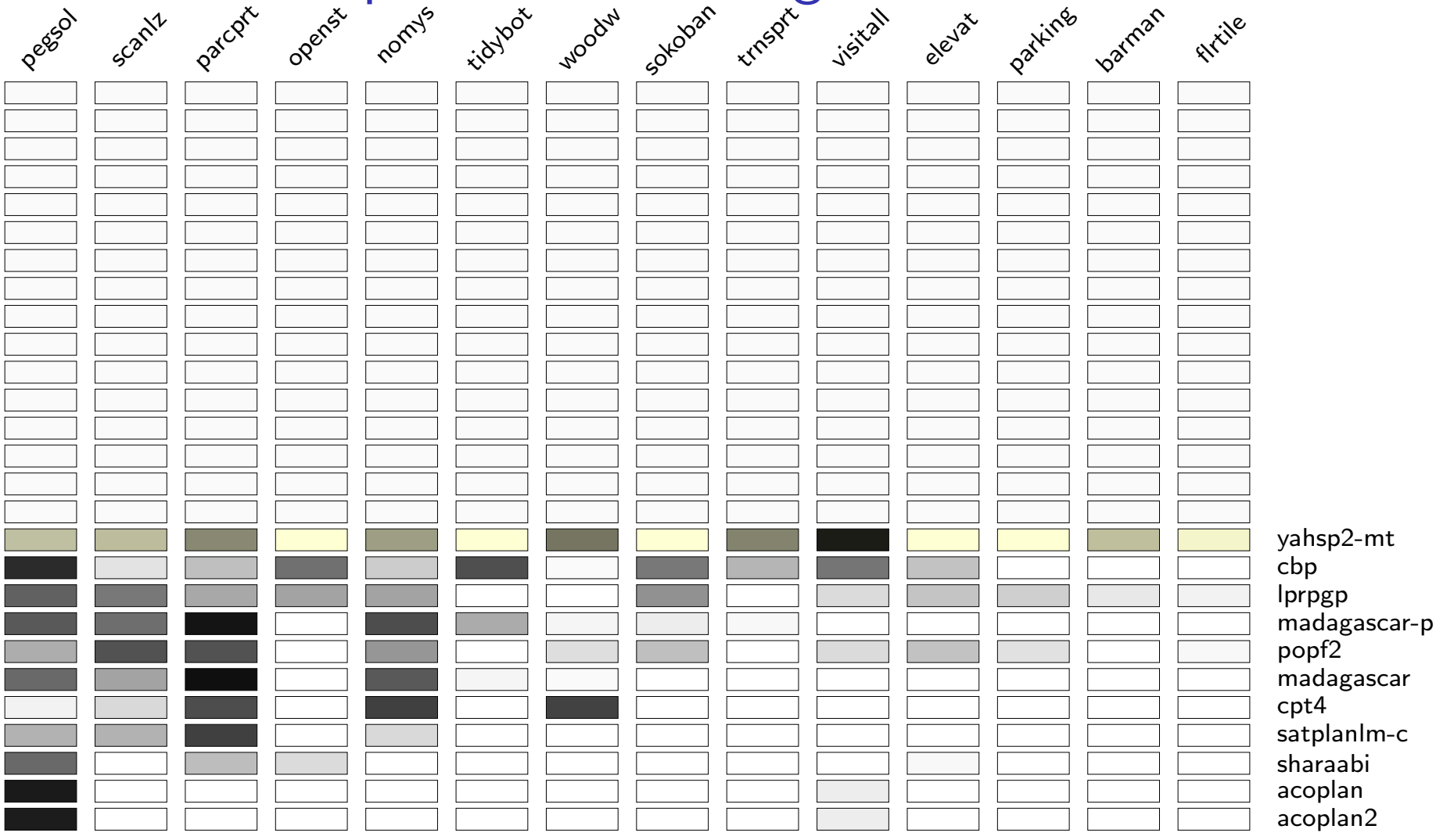
Sequential Satisficing track: Results



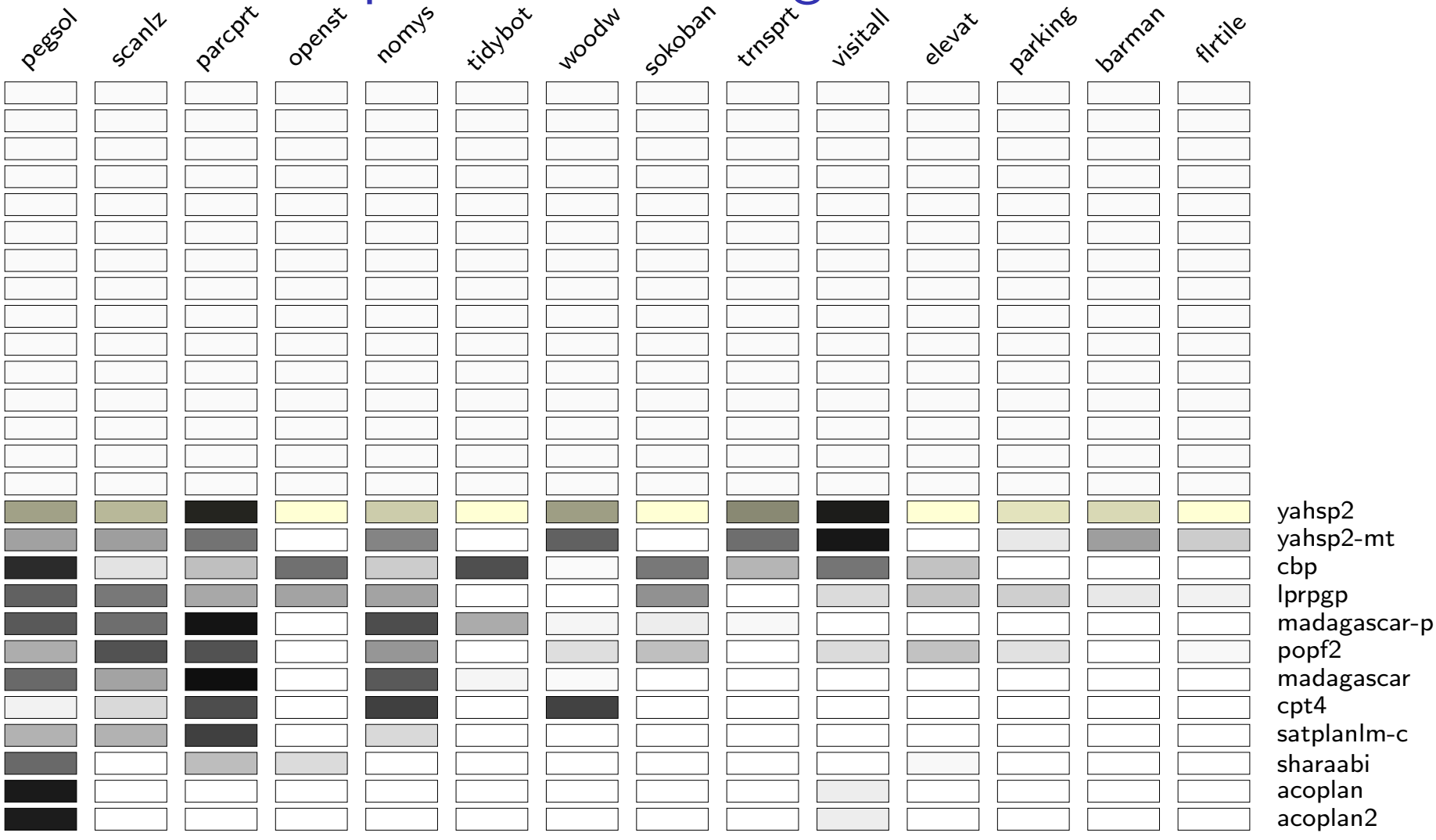
Sequential Satisficing track: Results



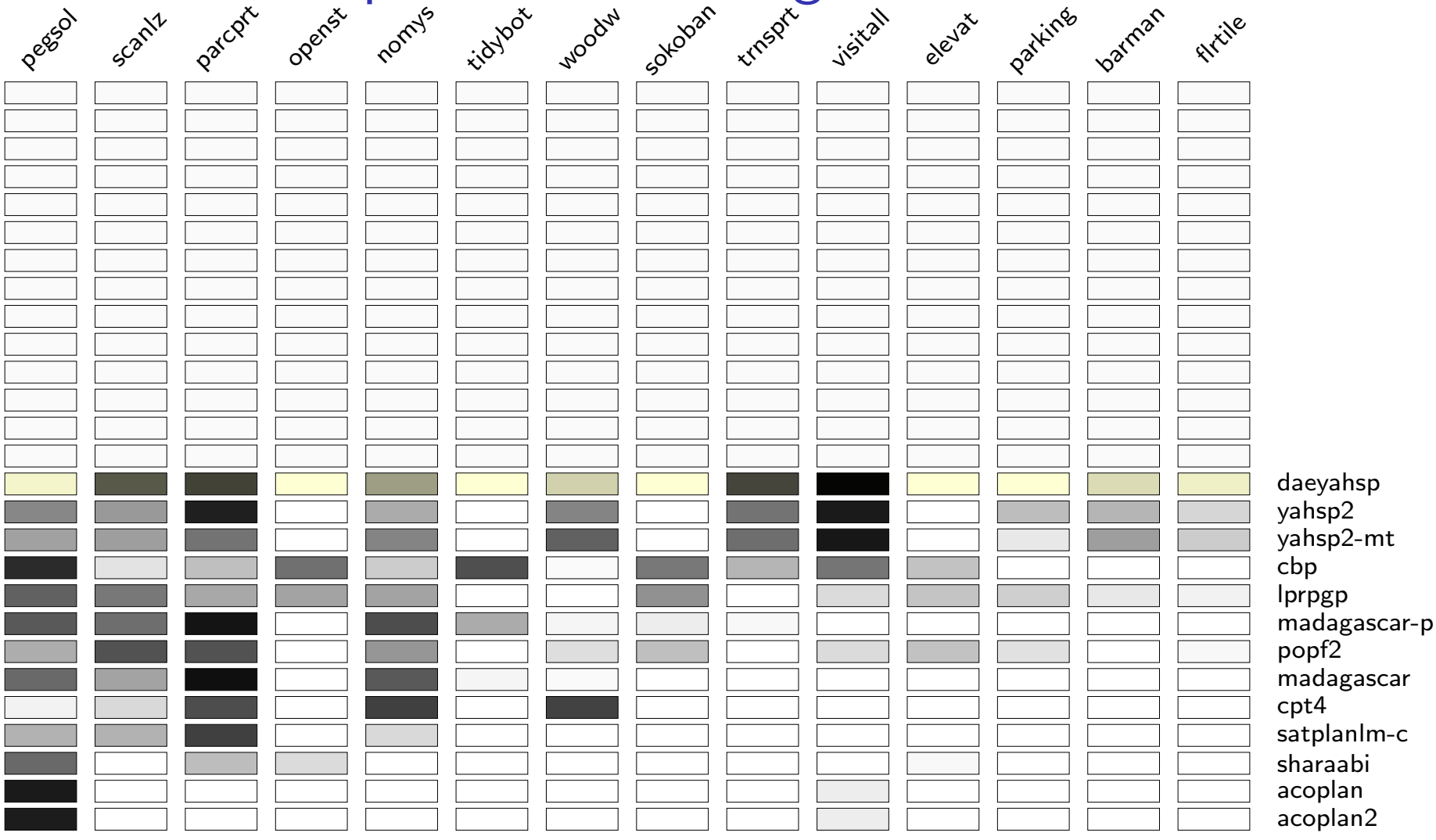
Sequential Satisficing track: Results



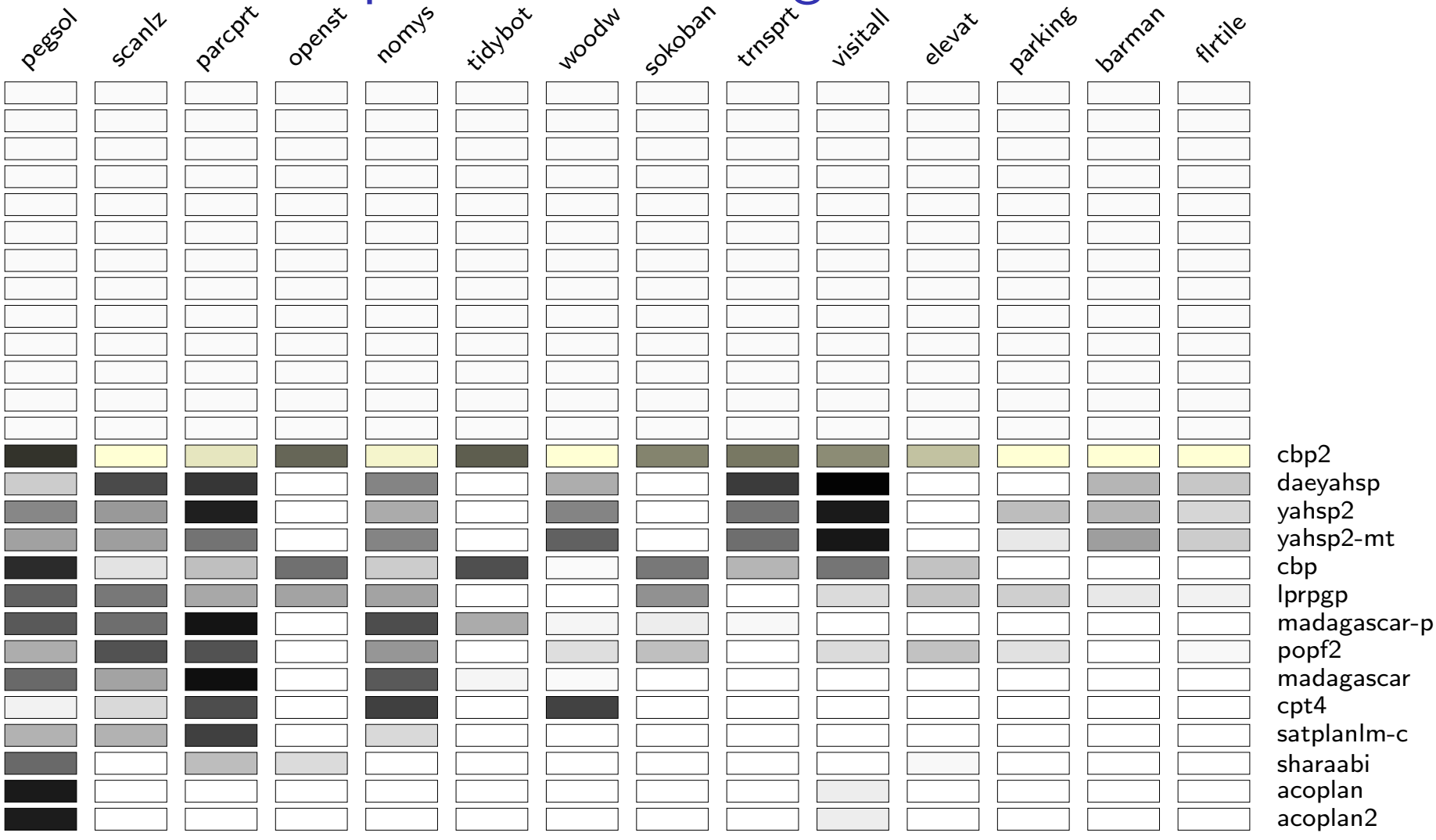
Sequential Satisficing track: Results



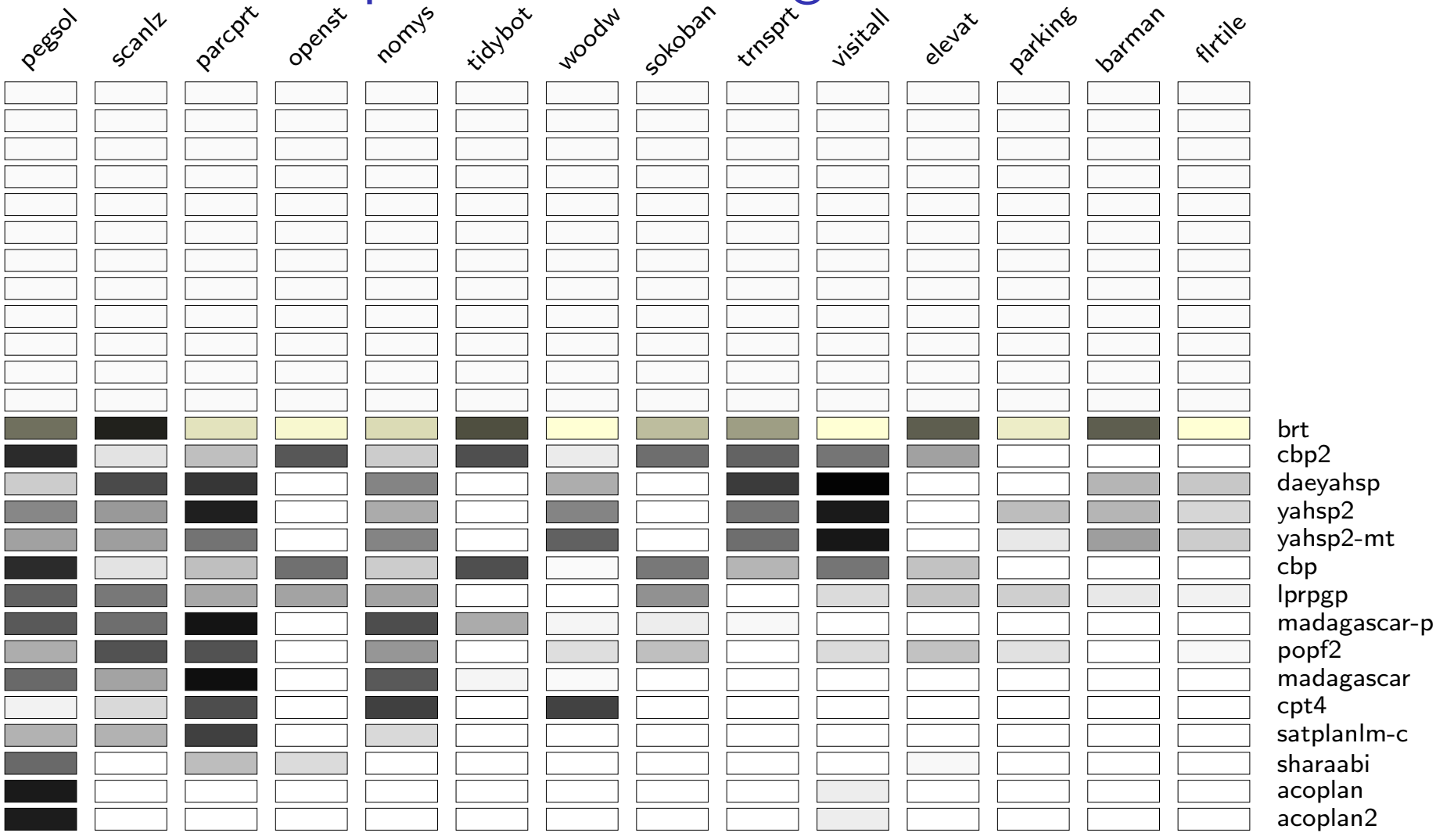
Sequential Satisficing track: Results



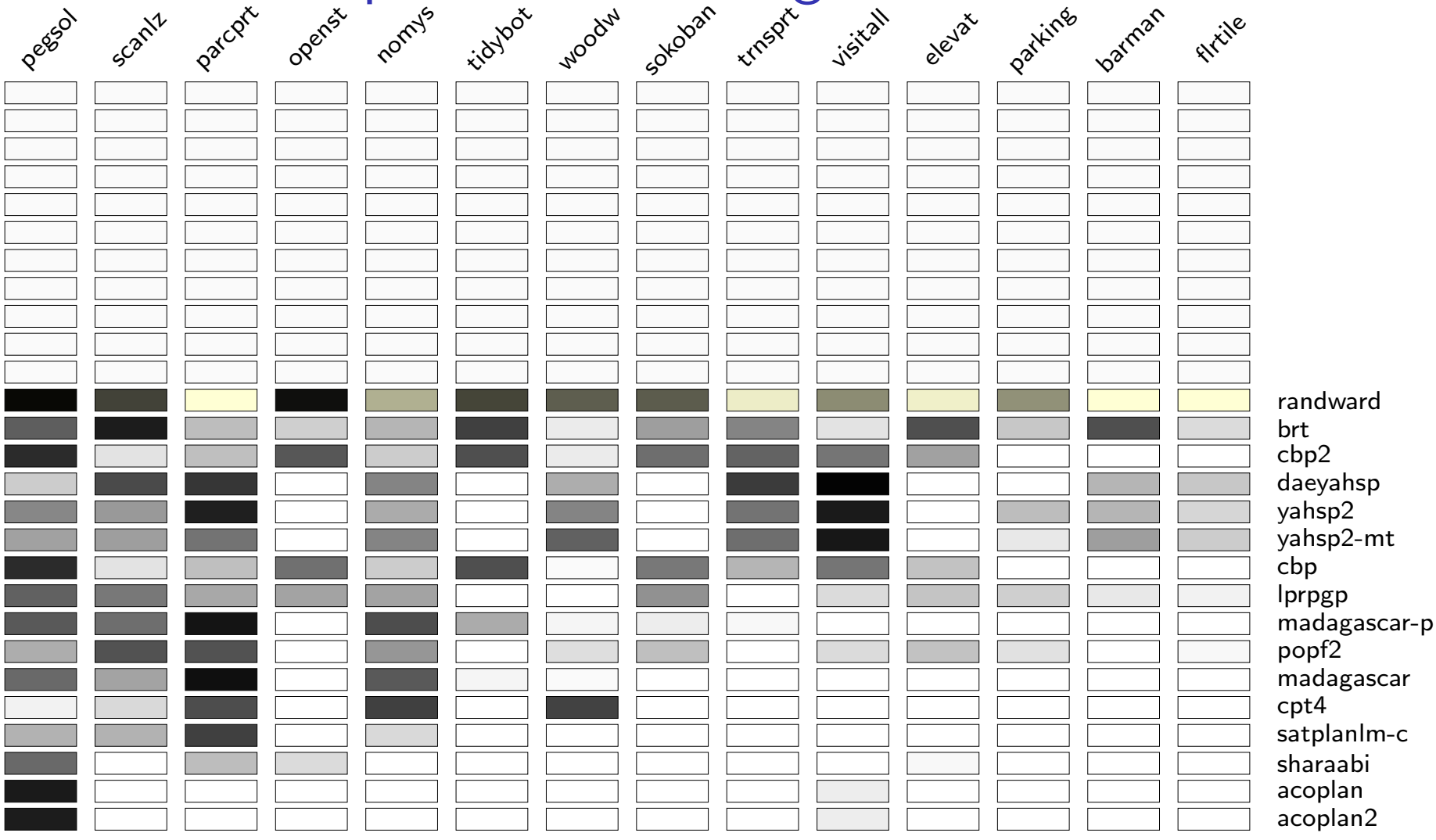
Sequential Satisficing track: Results



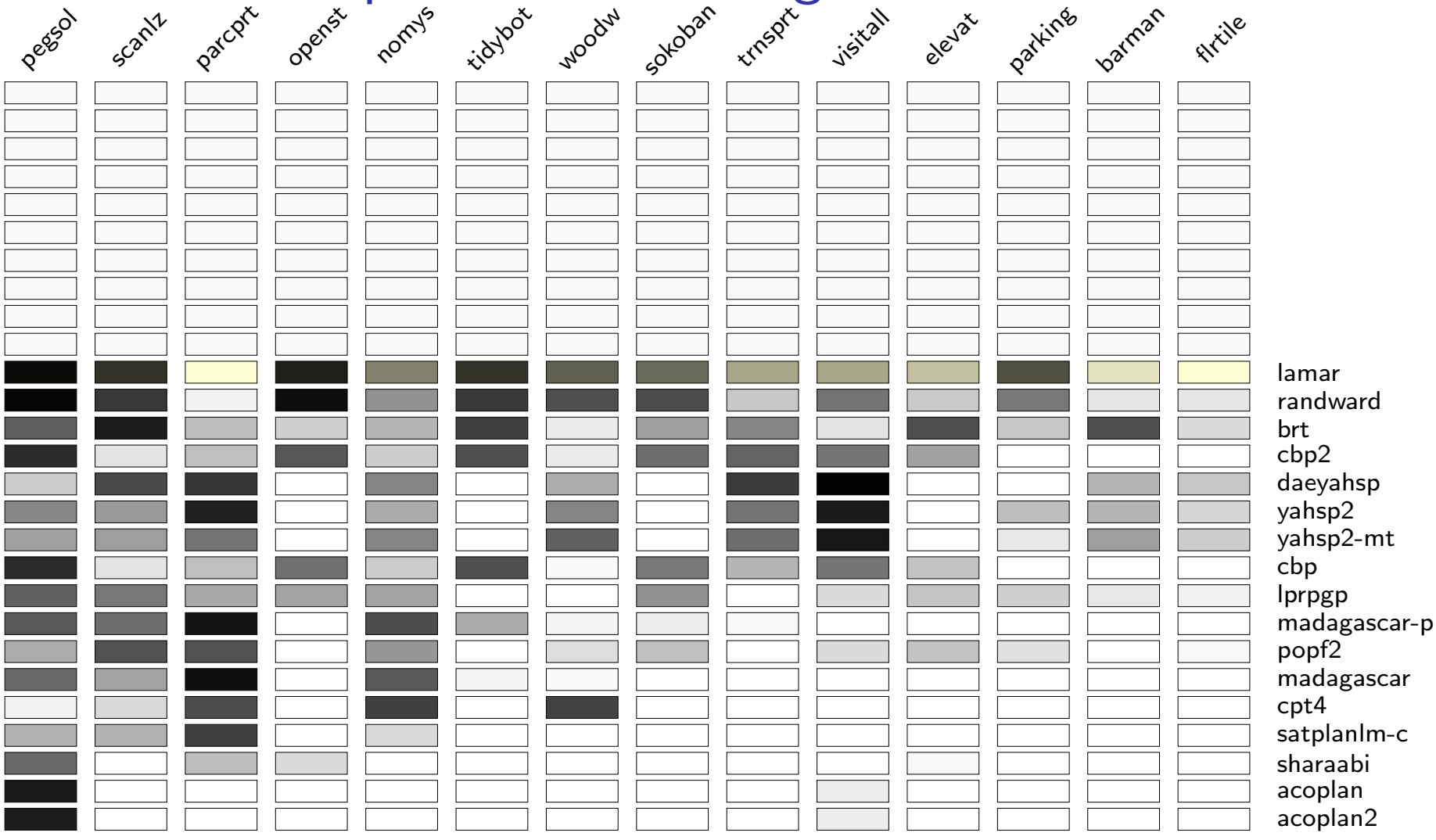
Sequential Satisficing track: Results



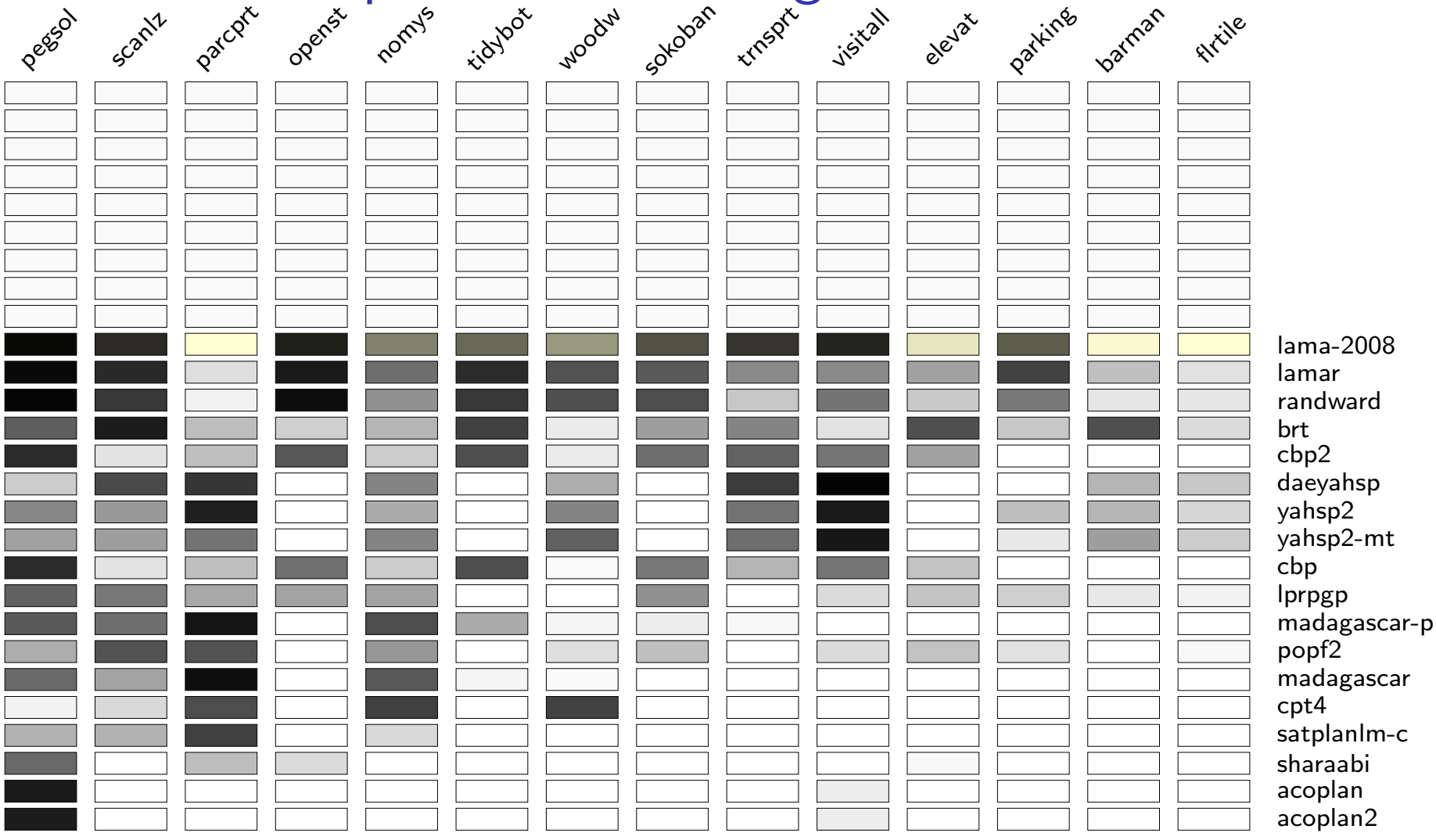
Sequential Satisficing track: Results



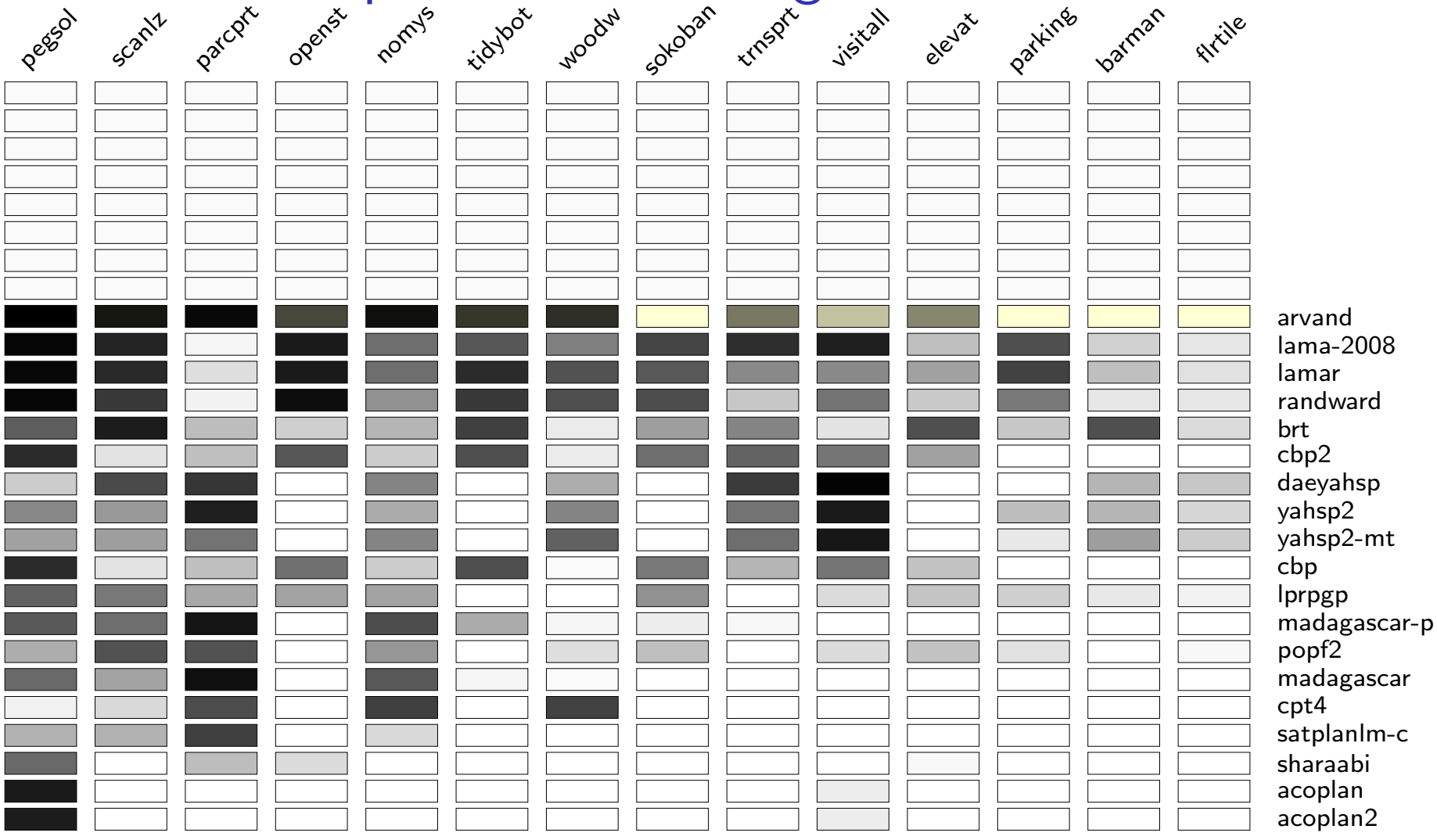
Sequential Satisficing track: Results



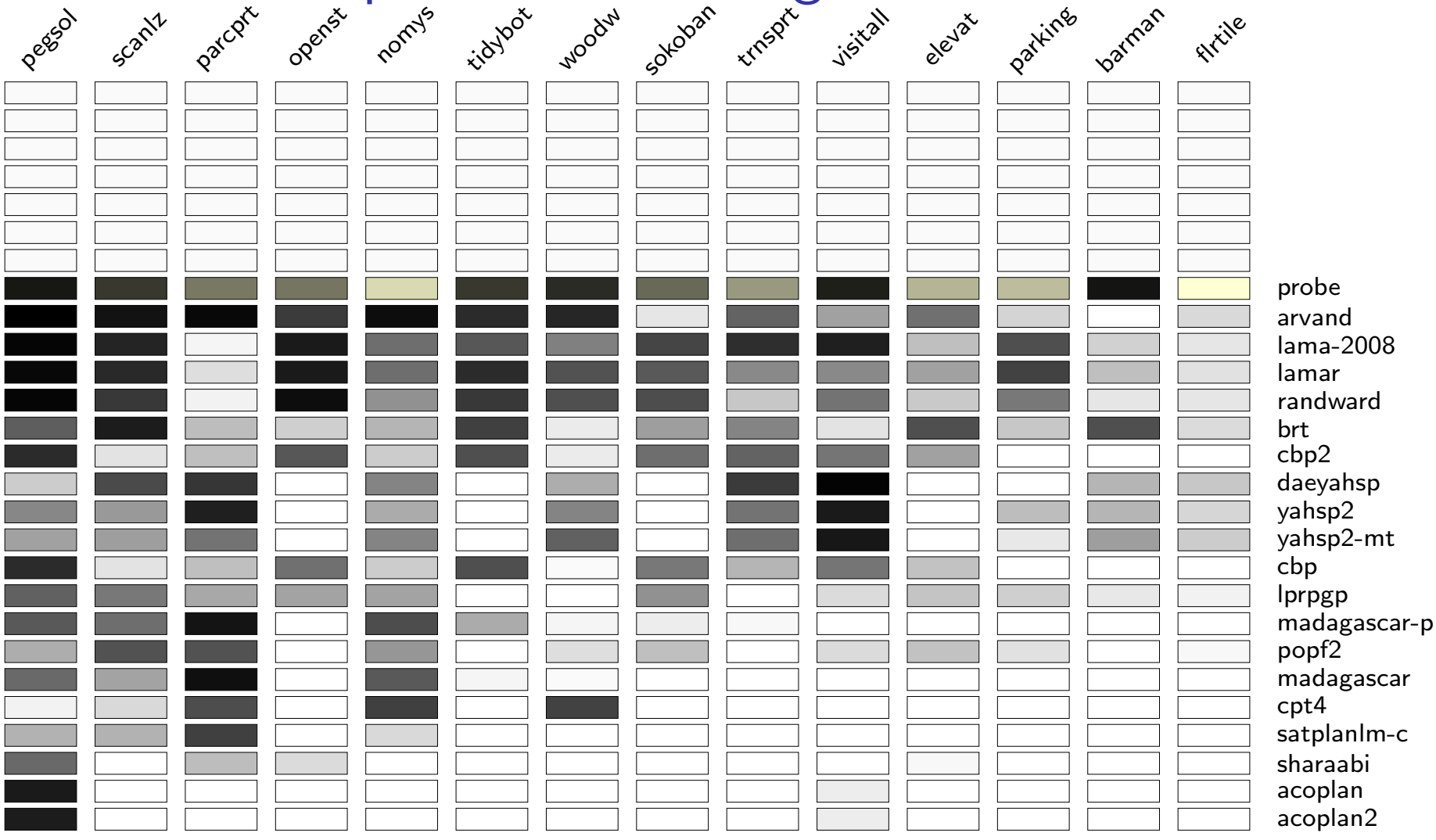
Sequential Satisficing track: Results



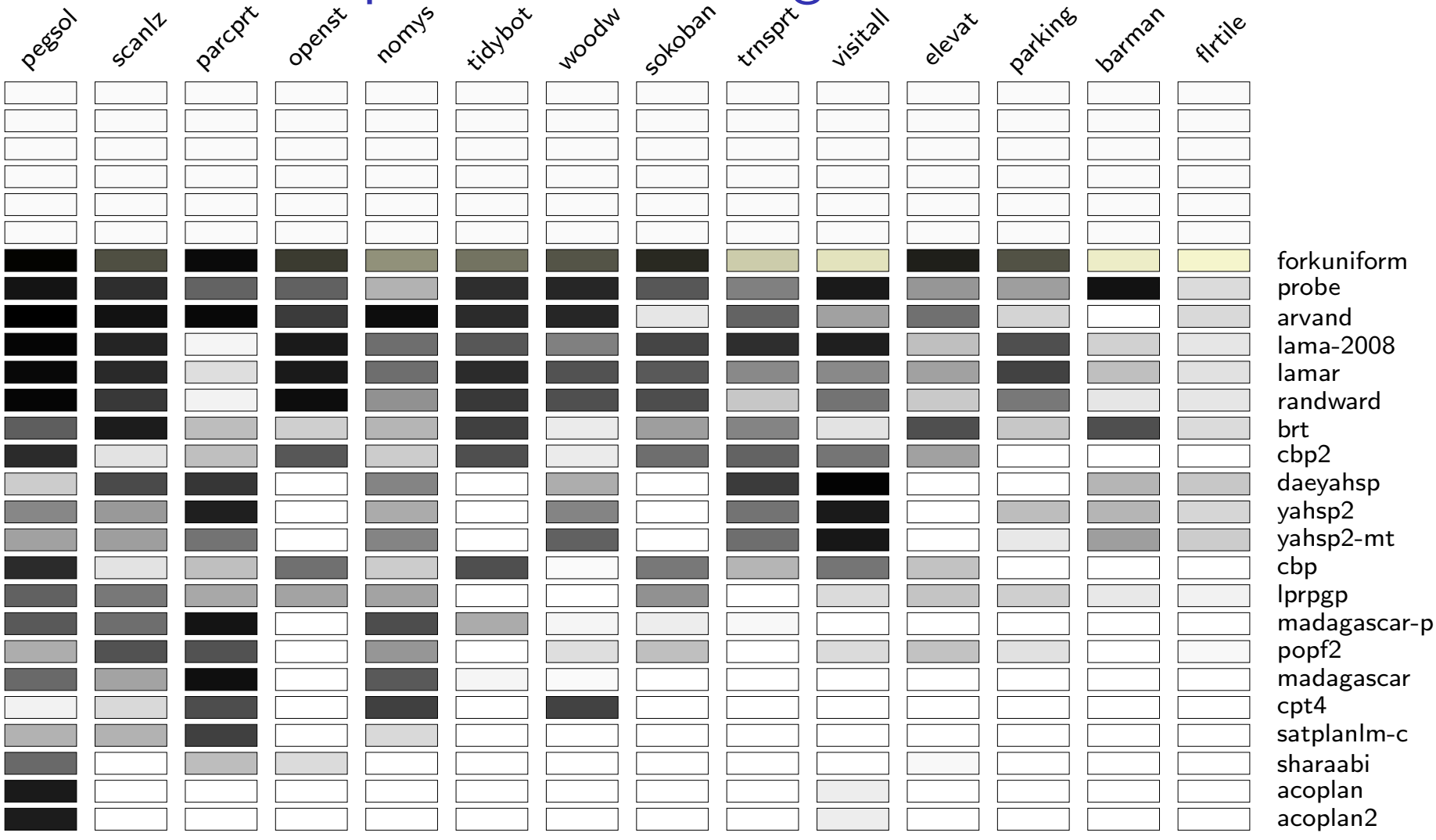
Sequential Satisficing track: Results



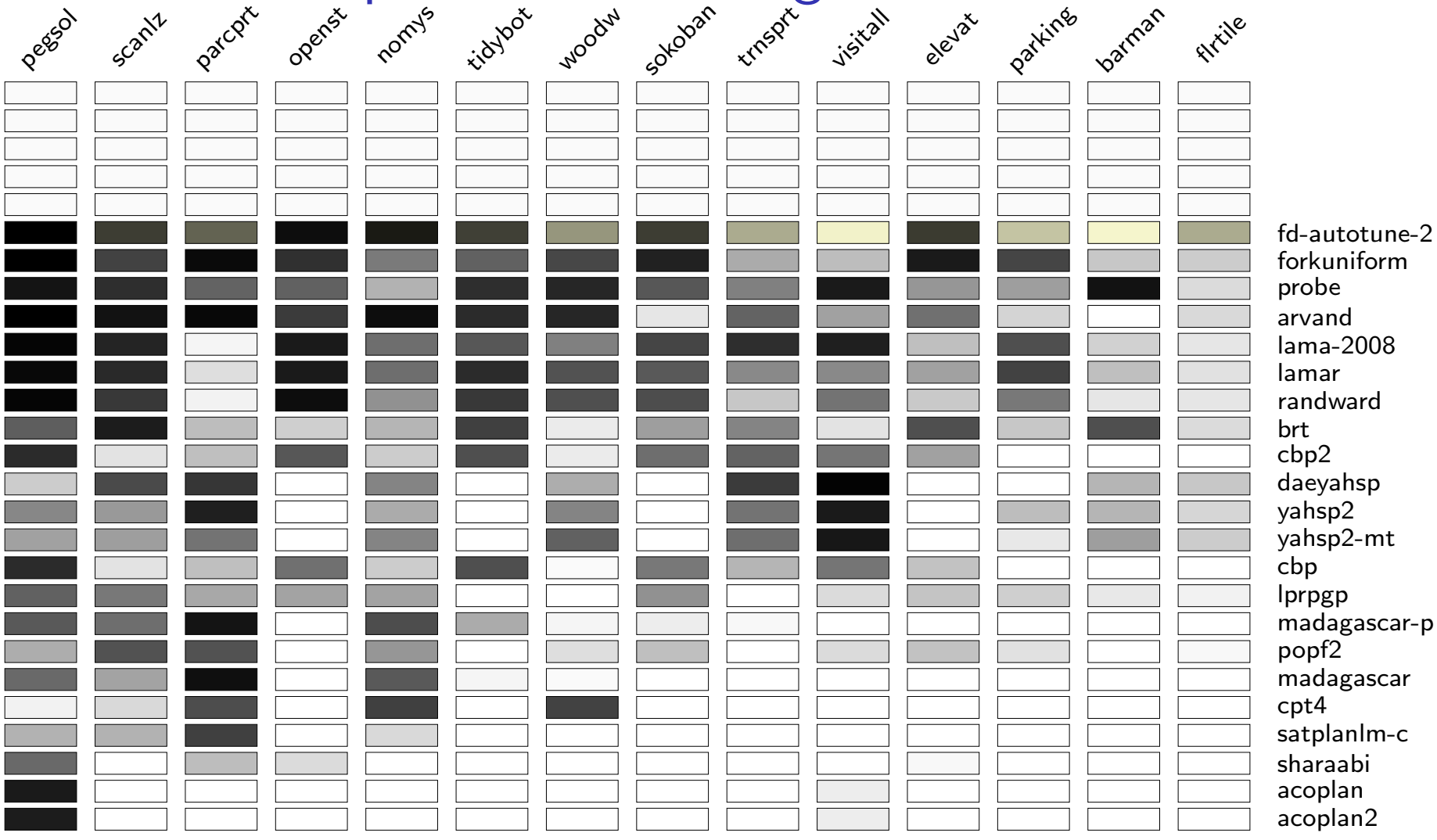
Sequential Satisficing track: Results



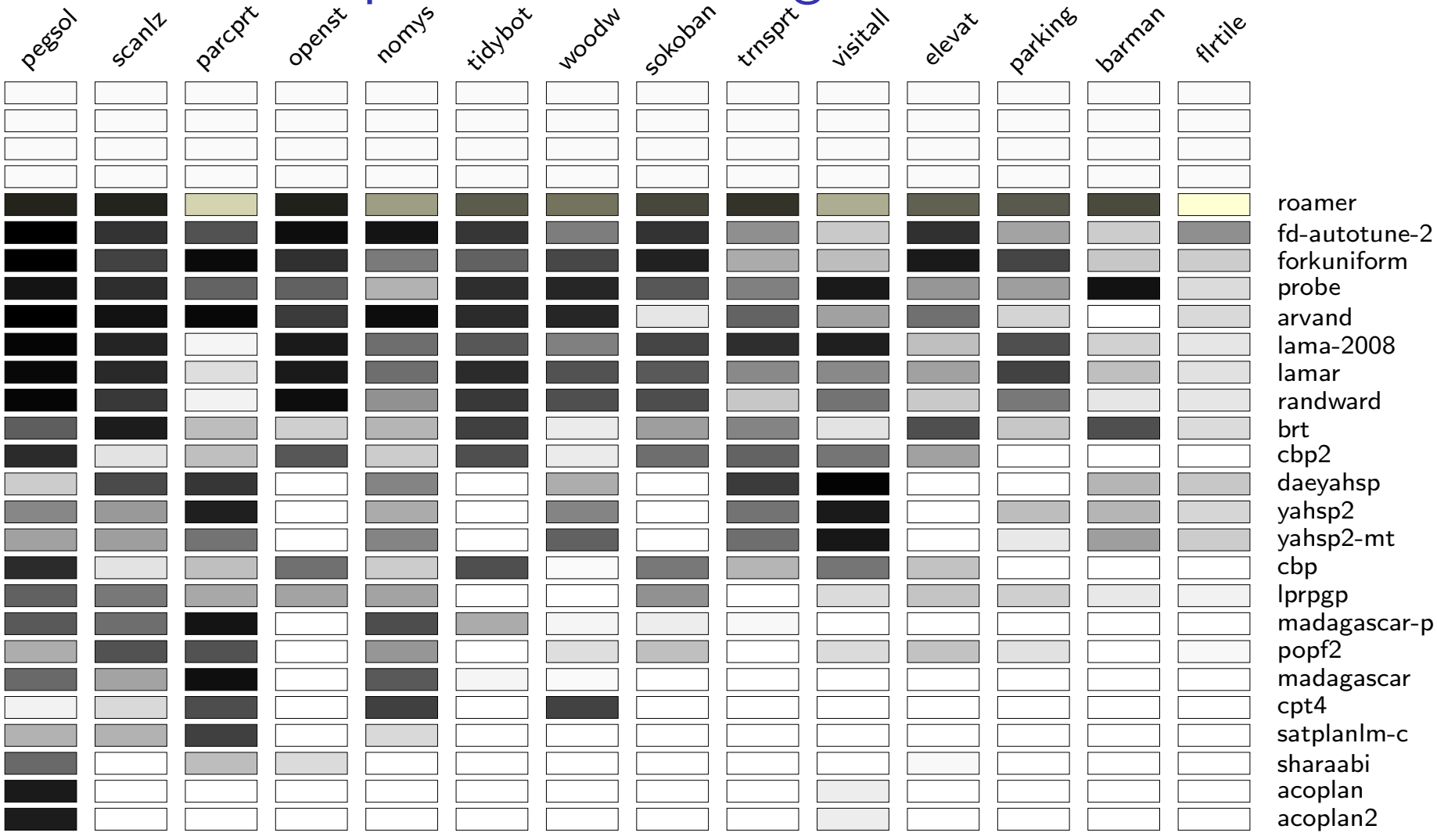
Sequential Satisficing track: Results



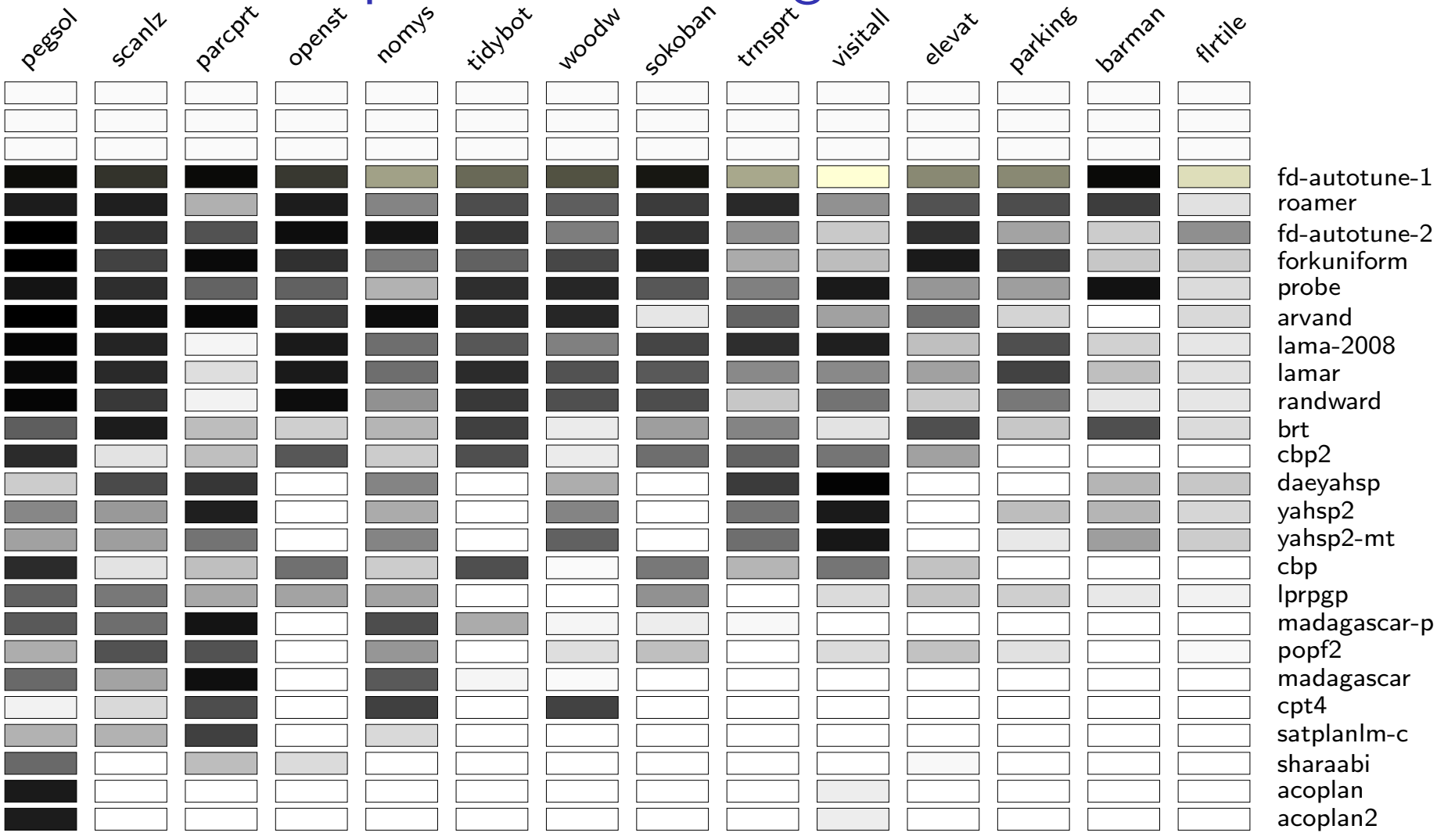
Sequential Satisficing track: Results



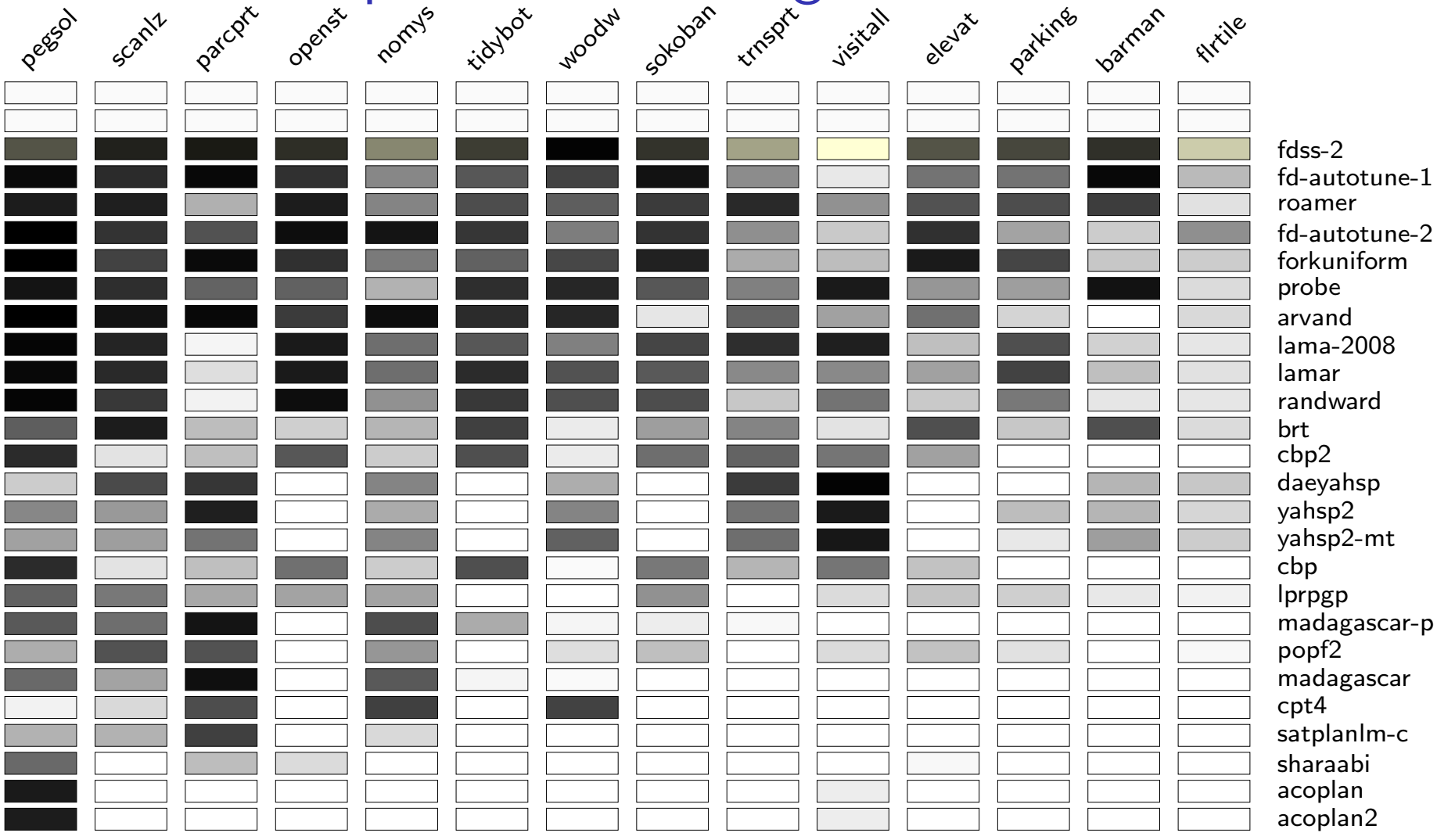
Sequential Satisficing track: Results



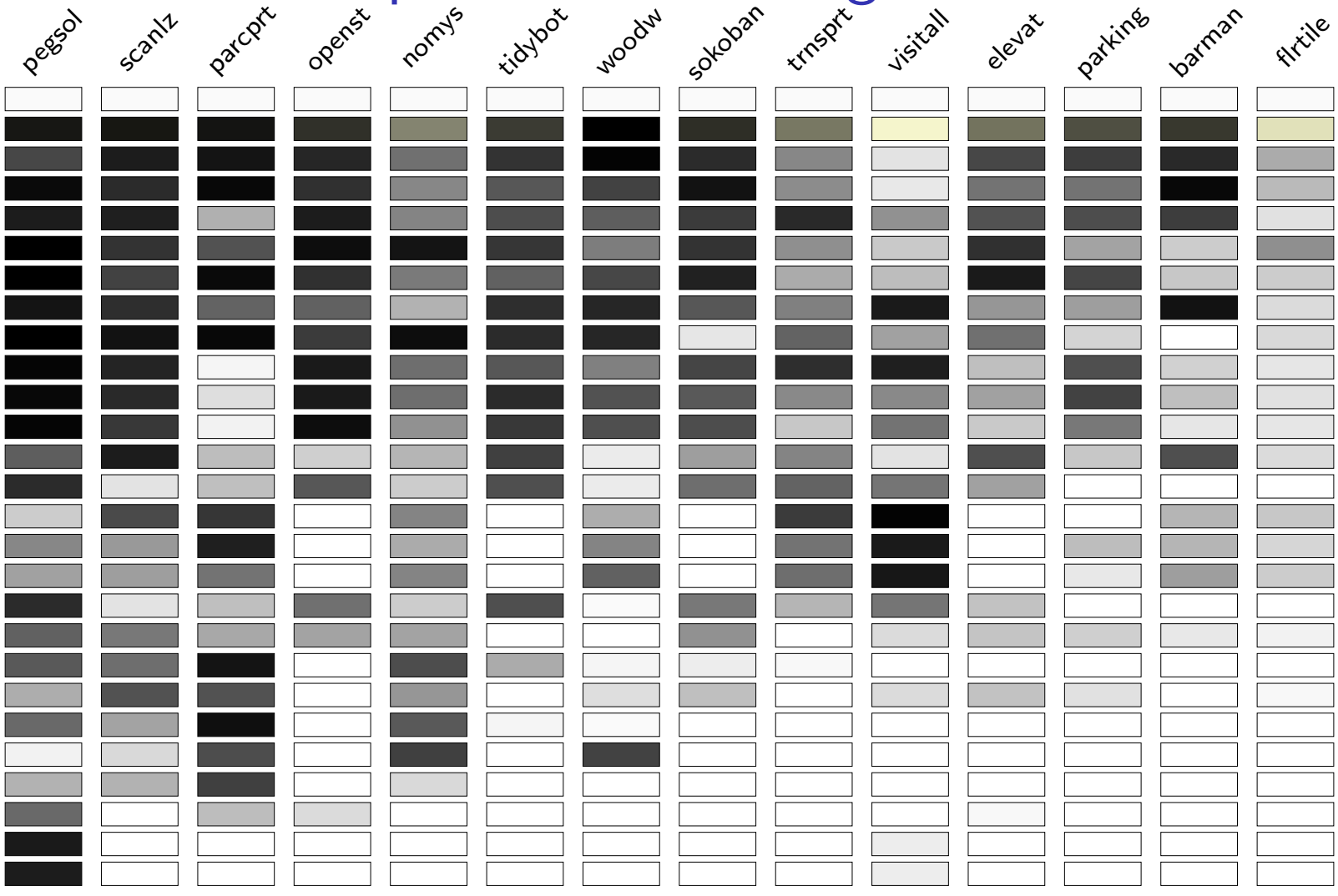
Sequential Satisficing track: Results



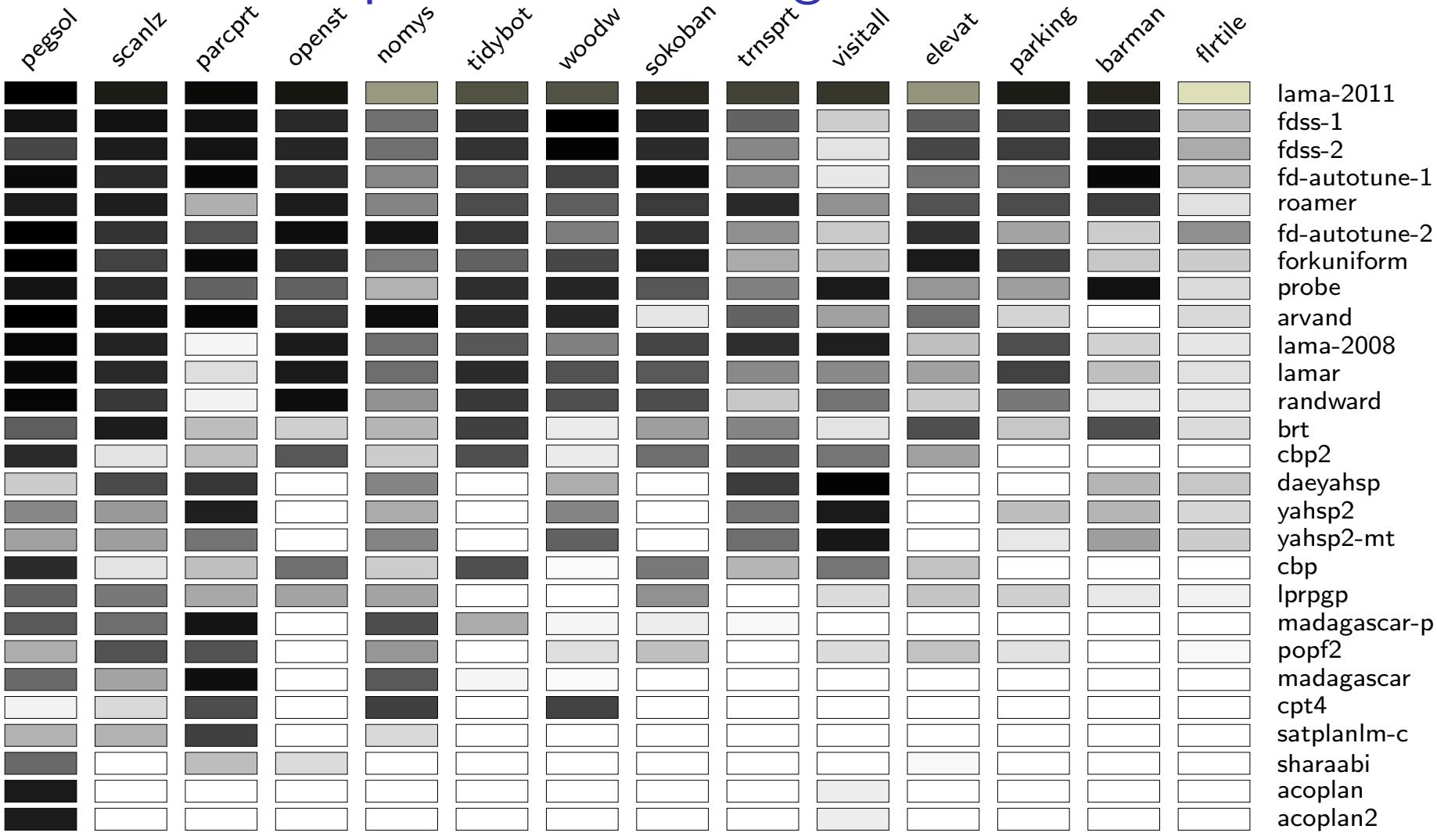
Sequential Satisficing track: Results



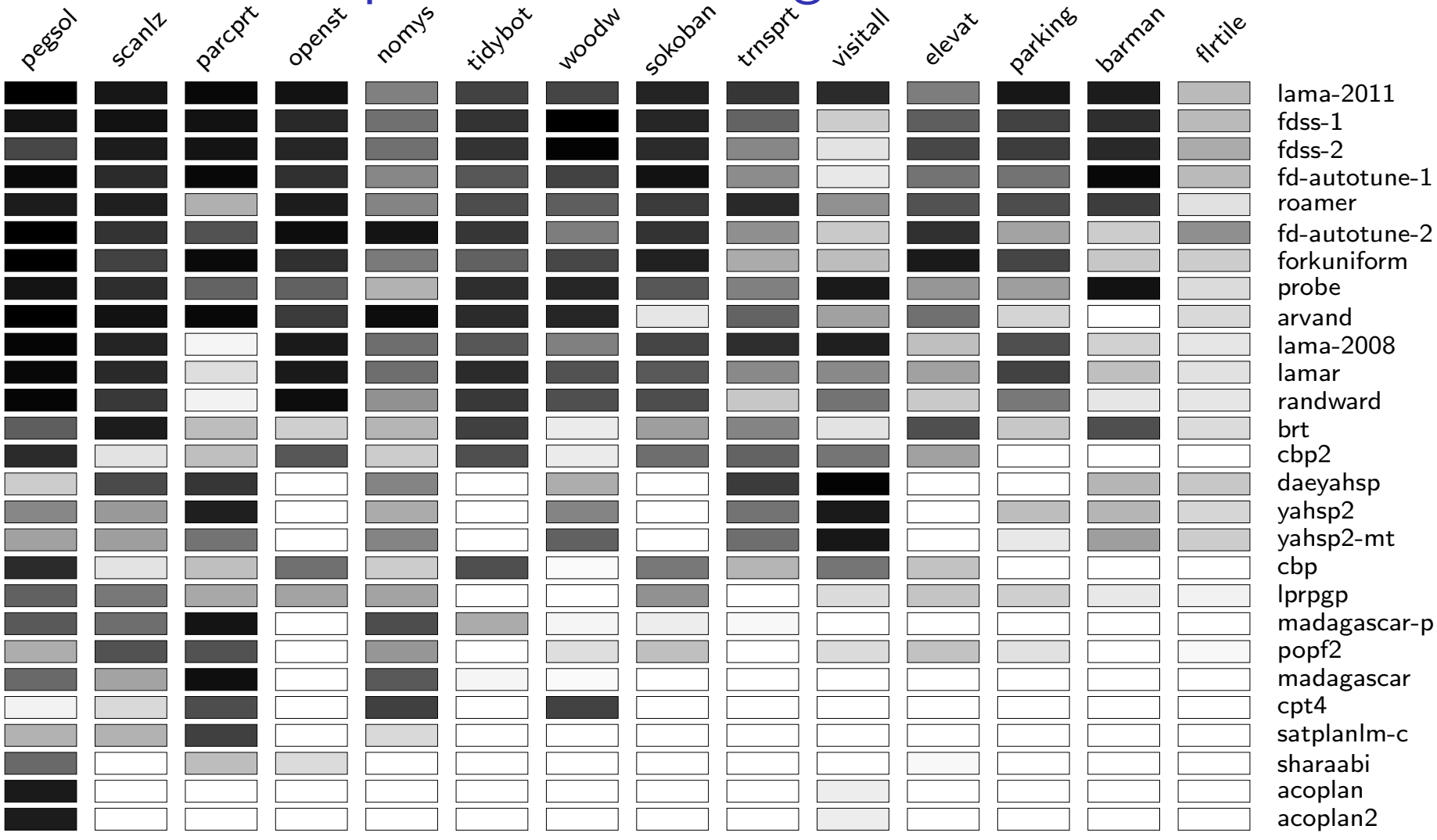
Sequential Satisficing track: Results



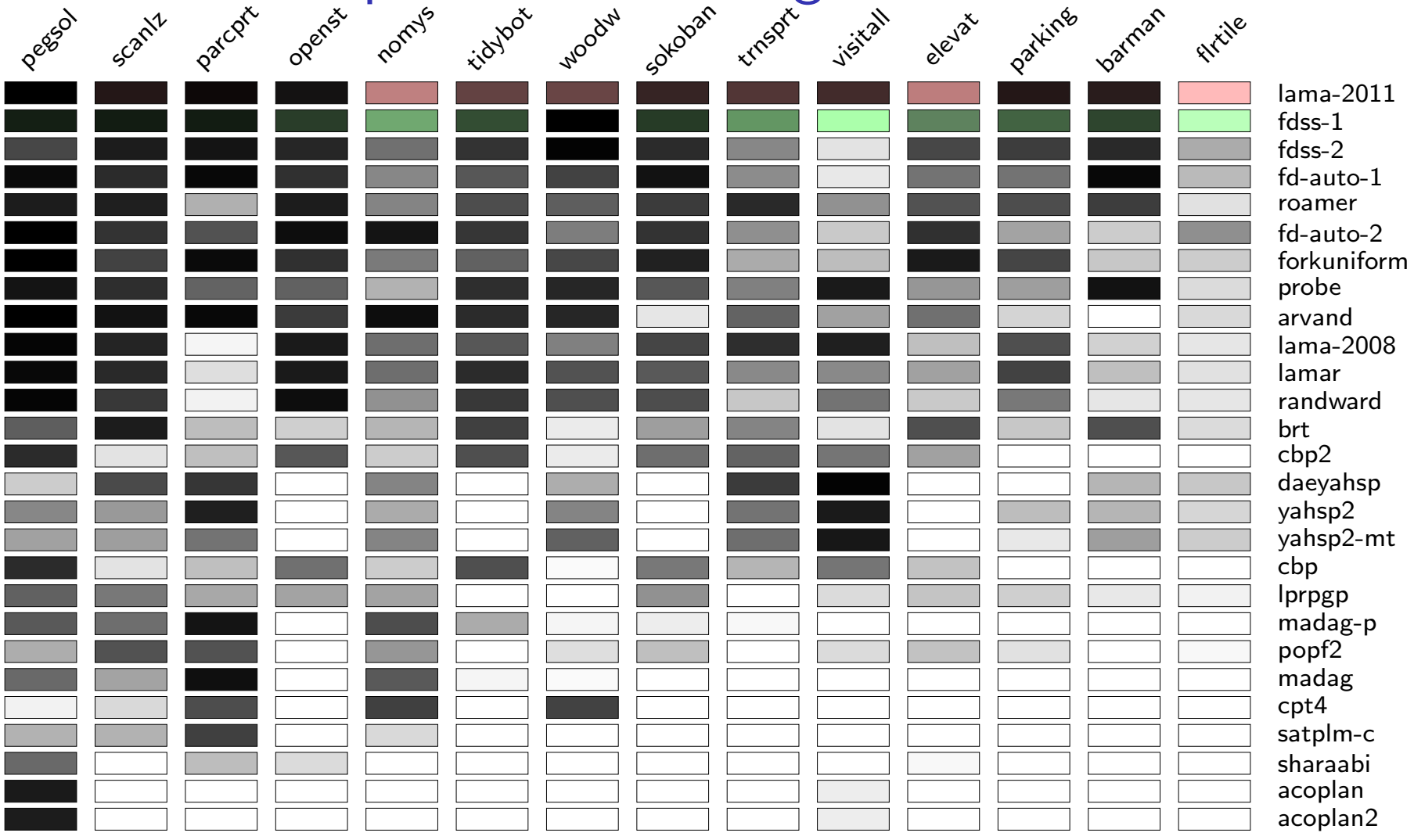
Sequential Satisficing track: Results



Sequential Satisficing track: Results



Sequential Satisficing track: Results



Sequential Satisficing track: Awards

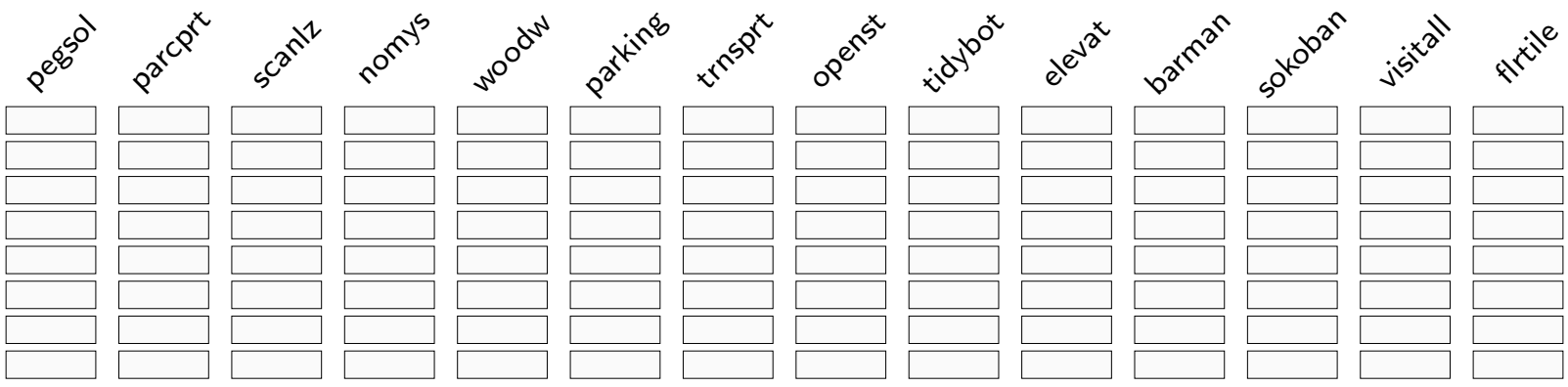
Winner

LAMA 2011: Silvia Richter, Matthias Westphal, Malte Helmert, Gabi Röger

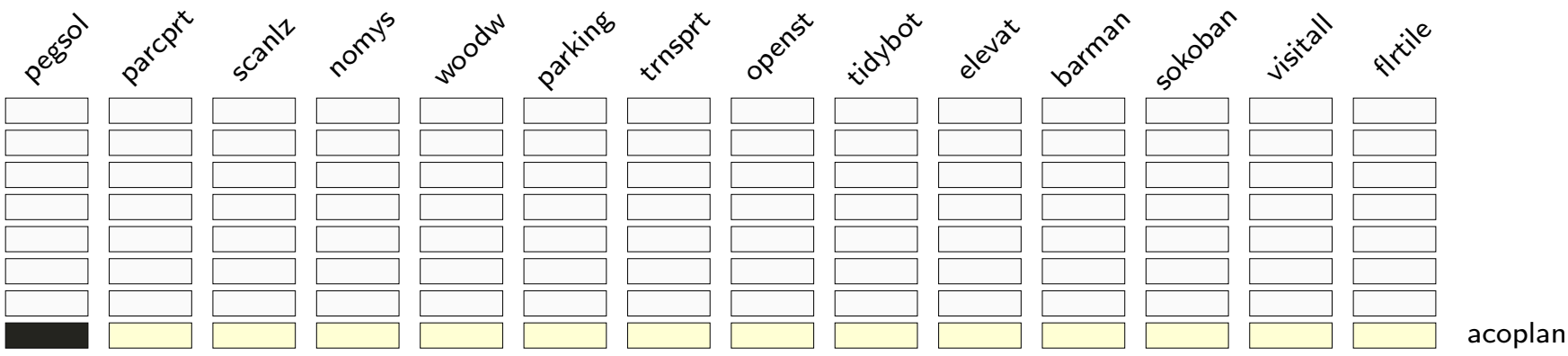
Runner-up

Fast Downward Stone Soup-1: Malte Helmert, Erez Karpas, Silvia Richter, Gabi Röger, Jendrik Seipp

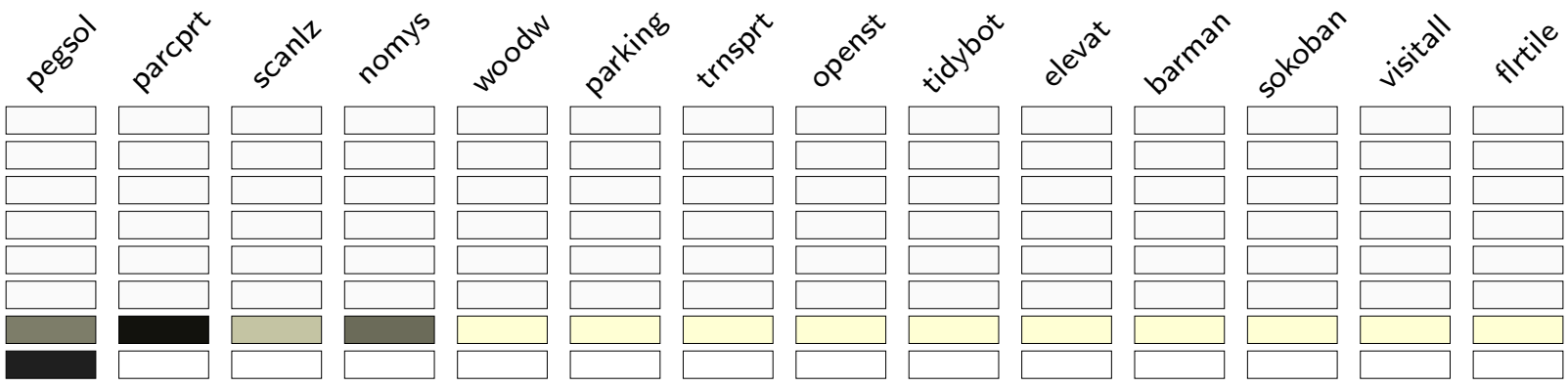
Sequential Multi-core track: Results



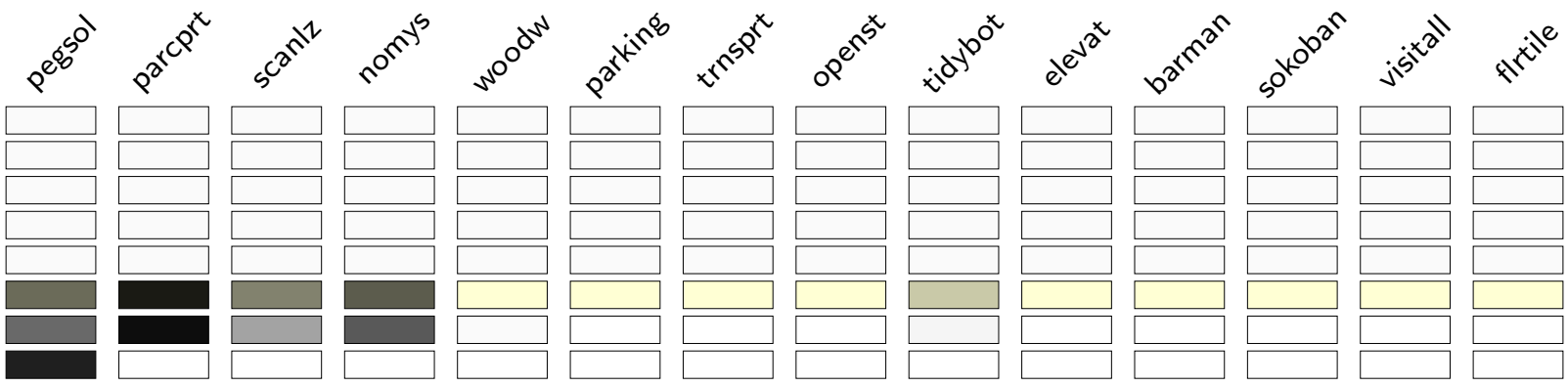
Sequential Multi-core track: Results



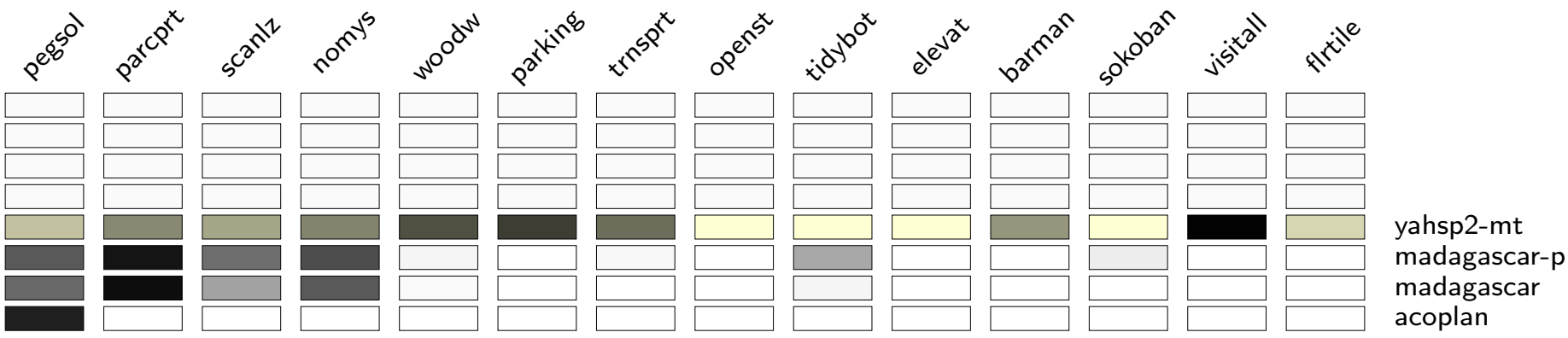
Sequential Multi-core track: Results



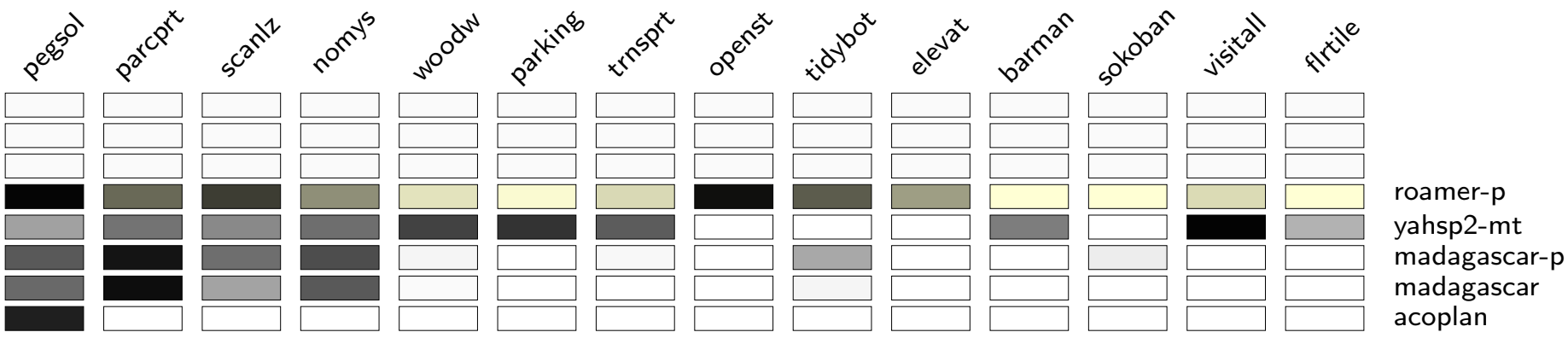
Sequential Multi-core track: Results



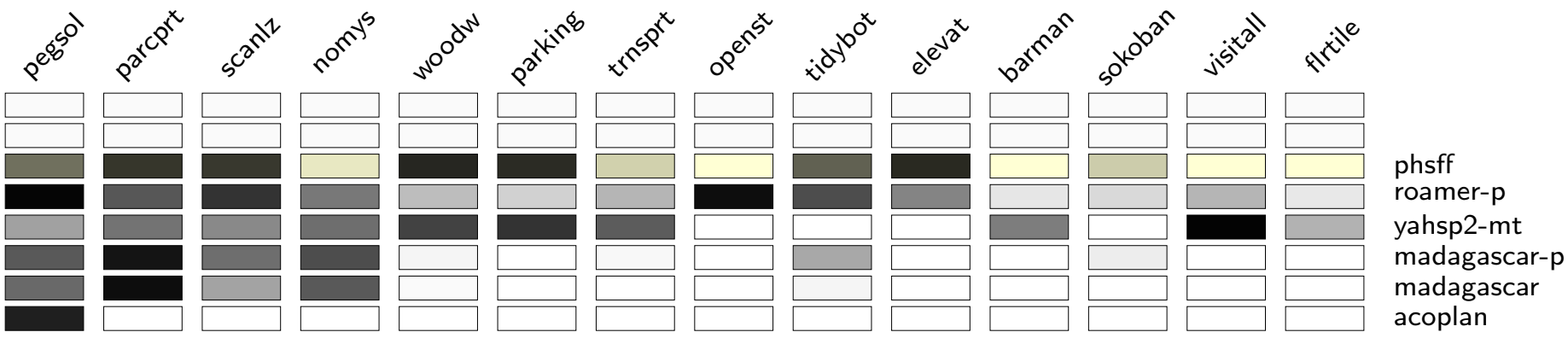
Sequential Multi-core track: Results



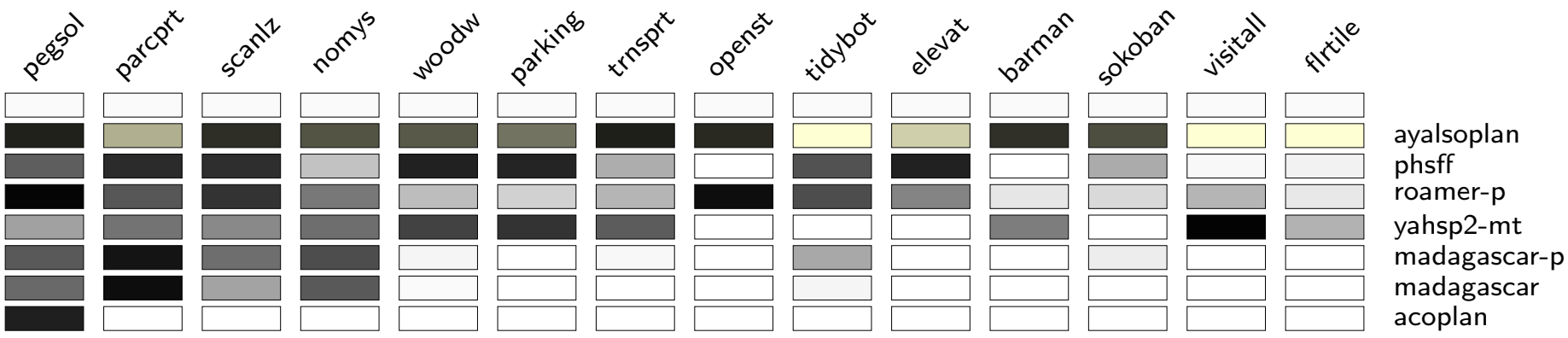
Sequential Multi-core track: Results



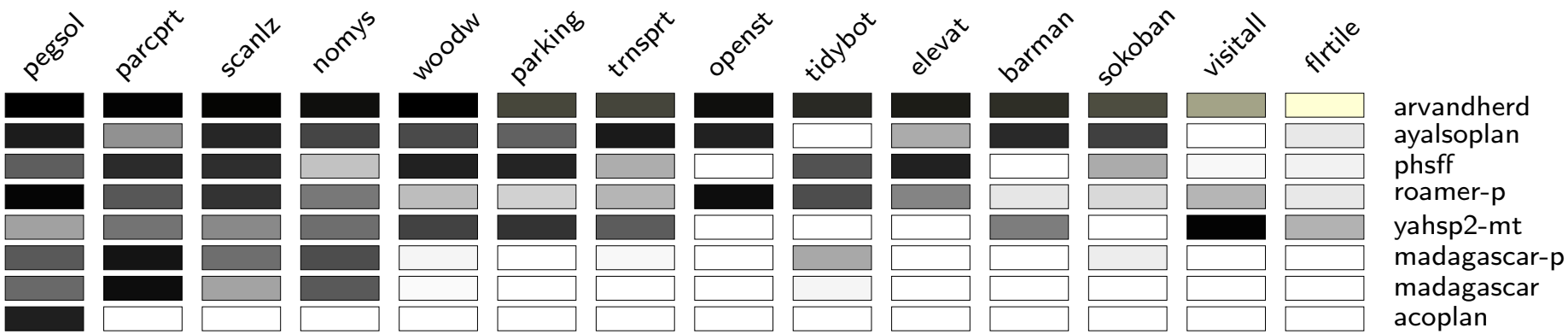
Sequential Multi-core track: Results



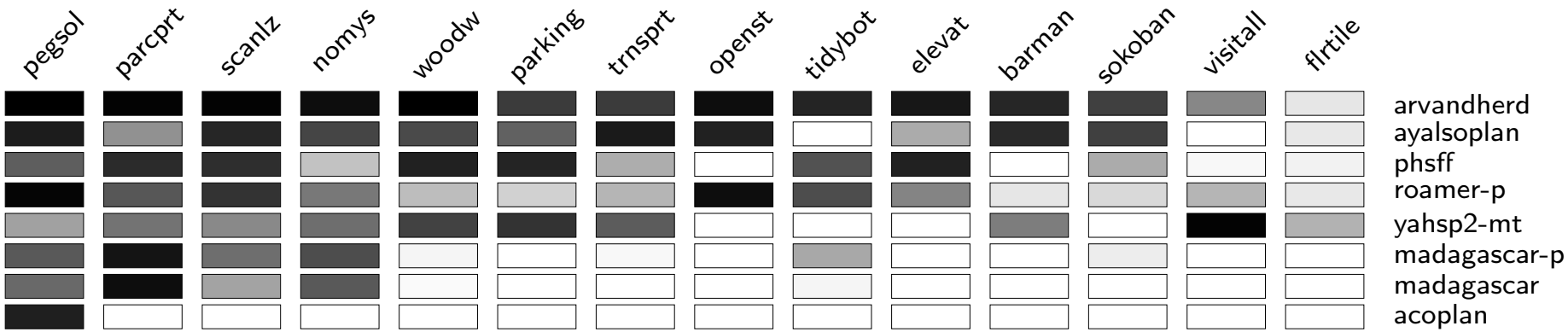
Sequential Multi-core track: Results



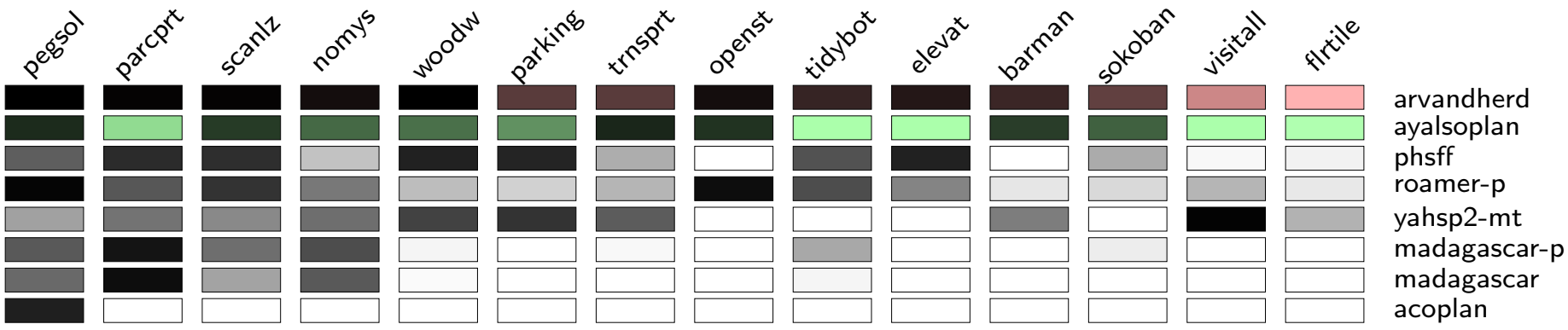
Sequential Multi-core track: Results



Sequential Multi-core track: Results



Sequential Multi-core track: Results



Sequential Multi-core track: Awards

Winner

Arvand Herd: Hootan Nakhost, Martin Mueller, Jonathan Schaeffer, Nathan Sturtevant, Richard Valenzano

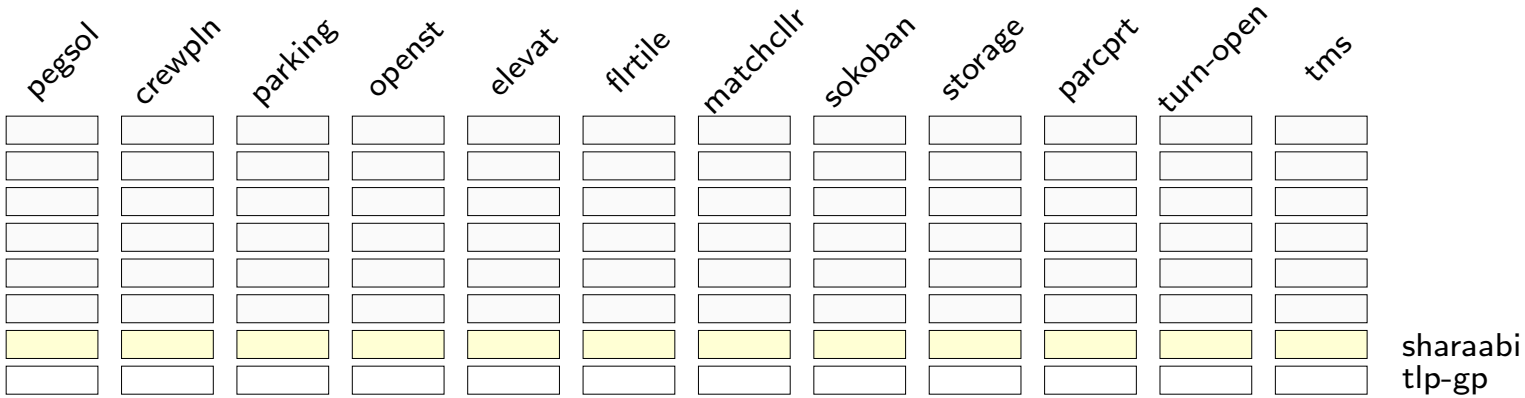
Runner-up

ay-Also-Plan Threaded: Juhan Ernits, Charles Gretton

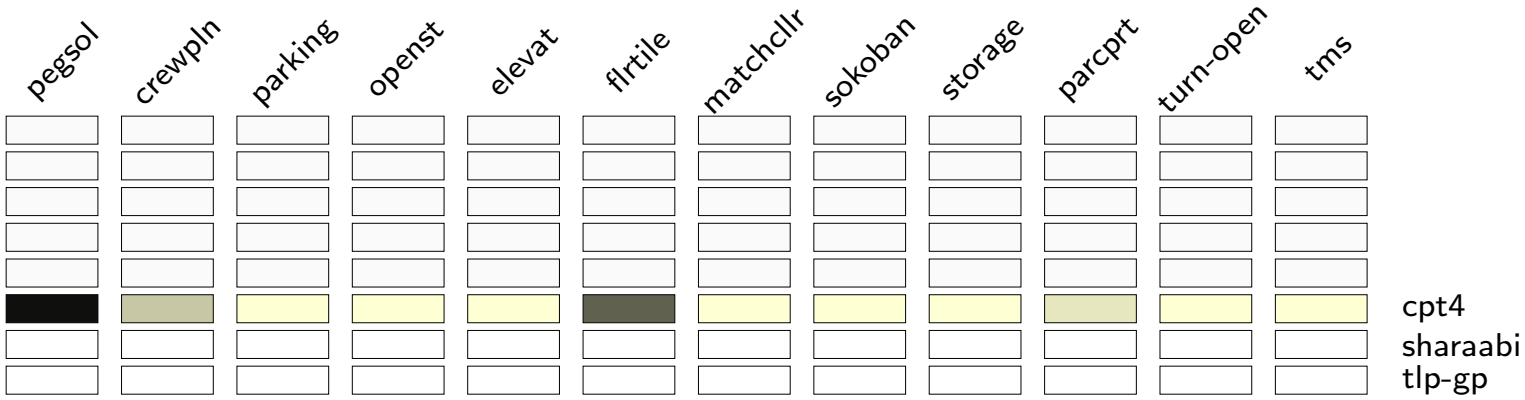
Temporal Satisficing track: Results

Temporal Satisficing track: Results

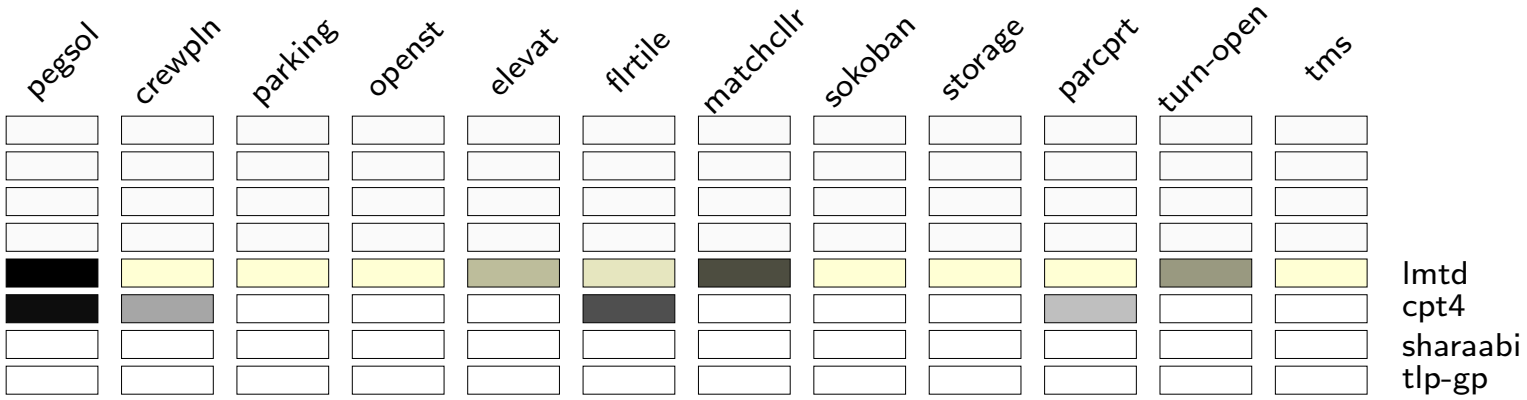
Temporal Satisficing track: Results



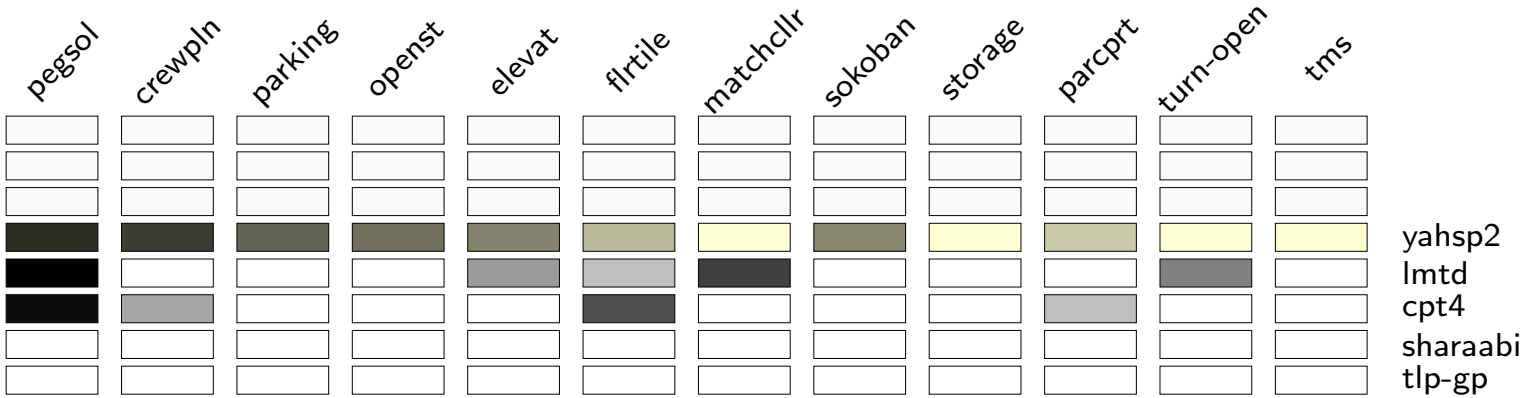
Temporal Satisficing track: Results



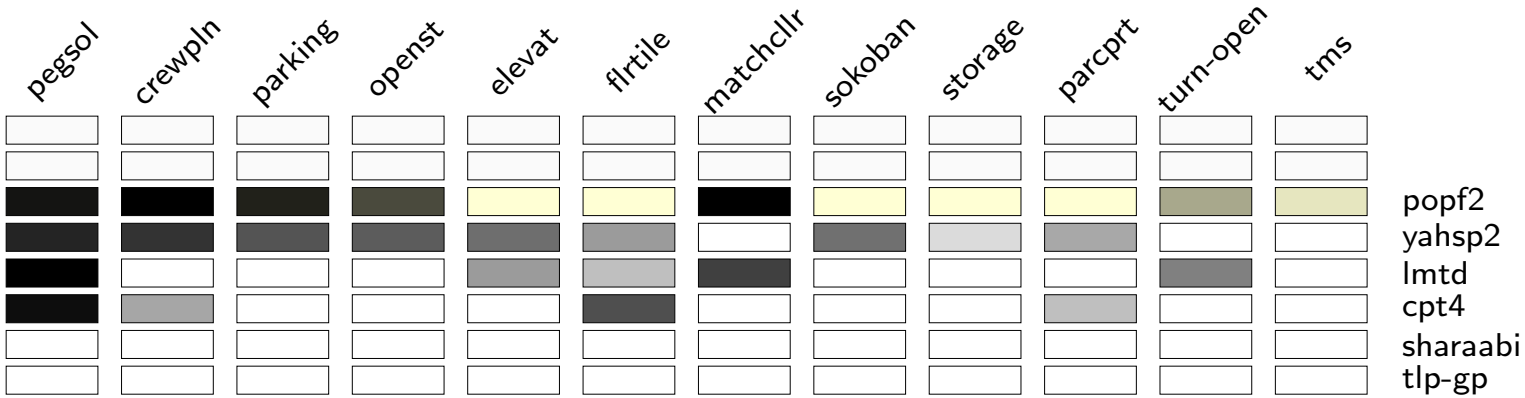
Temporal Satisficing track: Results



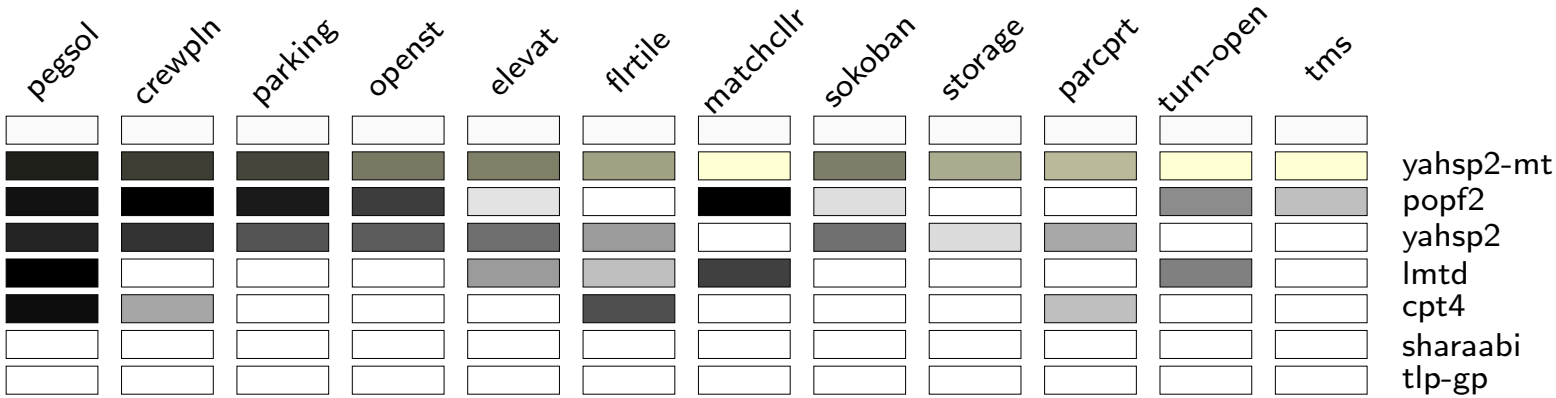
Temporal Satisficing track: Results



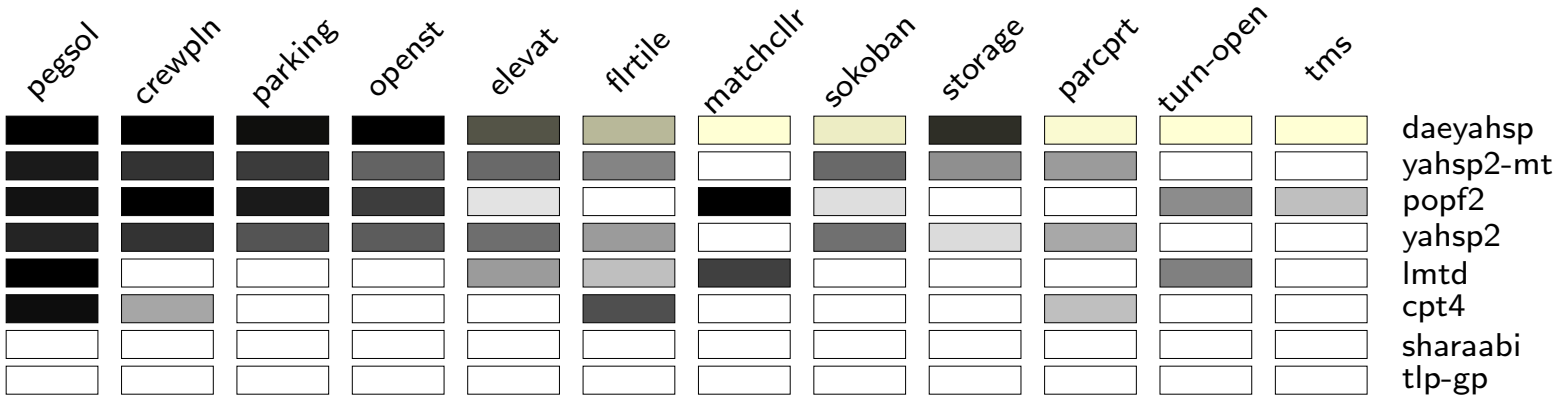
Temporal Satisficing track: Results



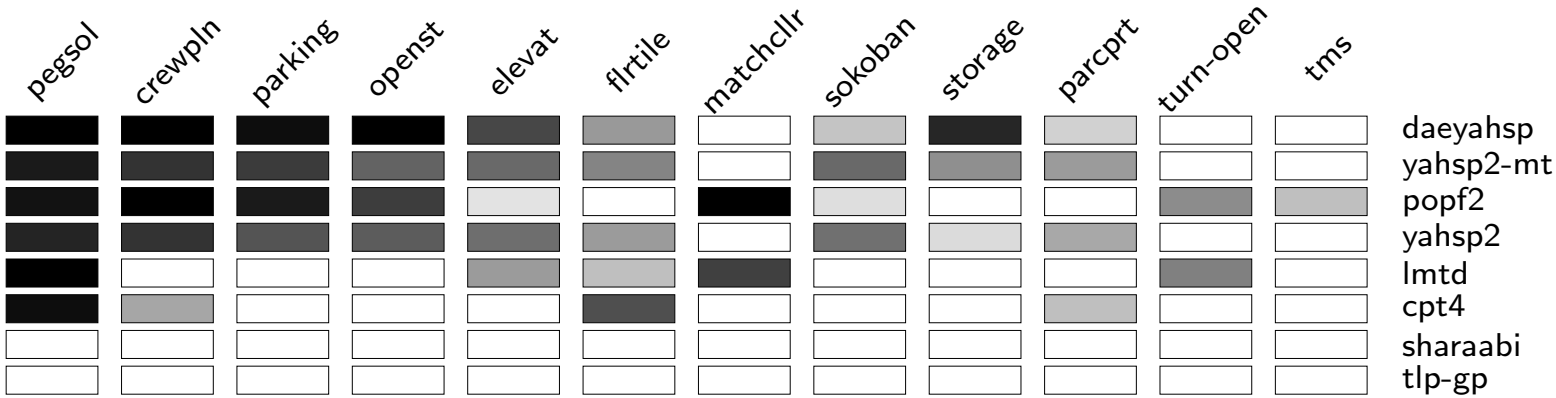
Temporal Satisficing track: Results



Temporal Satisficing track: Results



Temporal Satisficing track: Results



Temporal Satisficing track: Results

Temporal Satisficing track: Awards

Winner

DAEYAHSP: Johann Dréo, Marc Schoenauer, Pierre Savéant, Vincent Vidal

Runner-up ex-aequo

YAHSP2-MT: Vincent Vidal

Runner-up ex-aequo

POPF2: Amanda Coles, Andrew Coles, Maria Fox, Derek Long

Outline

Experiences with the competition

The current competition format (cont'ed from previous ones) allows:

- Better reflection of domain-independent planning: blind submission, source publication, papers publication
- Transparency: well-defined tracks, well-known scoring, public repositories, software available
- Continuity: allows better comparisons among editions of the IPC
- No baseline planner needed

Experiences with the software

Sharing code between editions of the IPC allows:

- Repeatibility and error-checking: giving away the same source code and the same data used by the organizers
- Extendability: improving code between editions of the IPC saves a lot of burden!
- Stability: maintaining (*Did I say supporting?*☺) the code used at the IPC

Experiences with the language

PDDL has reached a remarkable degree of complexity:

- Most planners do not fully supported (if any!)
- The intersection of the features supported is:
  - STRIPS (+ fluents) in the sequential tracks
  - STRIPS + fluents in the temporal tracks —often no numerical preconditions allowed!

Products of the IPC 2011

With the competition we disseminate:

- Data to be studied by the community:
  - $\sim$ 35 Gb of both raw and validated data
  - $\sim$ 100 Mb in the form of snapshots for making fast queries
  - $\sim$ 250 pdf pages and 20 different Excel wordsheets with different data and scores
- Software to be used both privately and (hopefully!) in future editions of the IPC
- Documentation on the software and entrants:
  - Software is documented both in pdf and html
  - All papers will be published as a booklet —with the consent of authors

Our two cents for the next IPC

- Organizing the IPC is an **enriching** experience
- It is a lot of work but:
  - We gained a lot of experience
  - Also, we already have tools and documentation
  - And some tradition
- Reuse planners and problems from the previous IPCs
- Extend and improve the current one and give everything away
- Strength its scientific character by collecting data and disseminating them
- While we could not do it (*sigh!*☹) relate to other competitions, more remarkably ICKEPS
- Do not allow another three years gap!

Thanks for attending!

Congratulations to the winners!