AIPS-2000 Planning Competition

Fahiem Bacchus
University of Toronto

Overview

- AIPS-98 featured the first competition.
  - 4 competitors in a STRIPS track, 2 in an ADL track
- This years competition:
  - 15 competitors
  - A fully automatic track with STRIPS & ADL domains
  - A hand-tailored track allowing domain dependent information.

Make Possible By

- Michael Ady, Winter-City Software, Edmonton, Alberta, Canada.
- John DiMarco, and the Department of Computer Science, University of Toronto, systems support staff.
- The other members of the organizing committee: Henry Kautz, David E. Smith, Derek Long, Hector Geffner, & Jana Koller
- The competitors who were willing to subject their work to a very public scrutiny.
- Franz Inc. For providing a free copy of Allegro Common lisp for Linux.

The Competitors

1. Blackbox

Yi-Cheng Huang
Bart Selman
Cornell University

Henry Kautz
AT&T Research

- Constructs a graphplan-graph, converts it into a Boolean satisfiability problems, and then attempts to solve the problem with various satisfiability engines.
- Competed in AIPS-98.
- Competed in the fully automated track.

The Competitors

2. MIPS

Stefan Edelkamp
Malte Helmert
University of Freiburg

- "Intelligent Model checking and Planning System"
- Uses BDDs compactly store and maintain sets of propositionally represented states, and a heuristic symbolic search engine as well as a heuristic single state search engine.
- Competed in the fully automated track.

The Competitors

3. System R

Fangzhen Lin
Hong Kong University of Science and Technology

- Based on a regression/progression algorithm like a sound version of the original STRIPS algorithm.
- Competed in both the fully automated and hand tailored tracks.
- In the hand tailored track it used domain specific information about (1) the ordering of subgoals; (2) pruning of unachievable goals; and (3) the way a subgoal is solved by regressing it to a new conjunctive goal.

The Competitors

4. FF

Joerg Hoffmann
Albert Ludwigs University

- FF (Fast-Forward) employs heuristic search like HSP, but extending the HSP heuristic to include information from GRAPHPLAN's plan extraction phase.
- Employs a local search strategy that combines Hill-climbing with systematic search.
- Competed the fully automated track.

The Competitors

5. HSP2

Hector Geffner
Blai Bonet
Universidad Simon Bolivar

- A heuristic-search planner that descends from the HSP planner that competed in the AIPS98 Contest.
- Planning instances are mapped into state-space search problems that are solved with heuristics extracted from the representation.
- HSP2 supports both forward and backward search and several heuristics, and uses a weighted A* search.
- Competed in the fully automated track.

The Competitors

6. IPP

Jana Koehler
Schindler Lifts Ltd.

Joerg Hoffmann
Michael Brenner
University of Freiburg

- IPP is based on searching GraphPlan planning graphs that have been extended to handle ADL actions.
- Essentially the same system as that entered in the AIPS-98 competition.
- Competed in the fully automated track.

The Competitors

7. PropPlan

Michael Fourman
University of Edinburgh

- PropPlan uses naive breadth-first state-space search but employs Ordered Binary Decision Diagrams to optimize its state space exploration.
- Operators are represented directly by efficient operations on BDDs.
- Like GraphPlan, the competition version of PropPlan utilizes forward chaining to establish a layered set of reachable states until the goal-set is reached, then backward chaining for plan extraction.
- Competed in the fully automated track.

The Competitors

8. TokenPlan

Yannick Meiller
Patrick Fabiani
ONERA - Center of Toulouse

- Based on the use of colored Petri nets which can encode mutex relations through the token's colors. Dependent on how tokens are propagated in the net a range of search techniques can be emulated.
- Competed in the fully automated track.

The Competitors

9. STAN

Maria Fox
Derek Long
University of Durham

- STAN employs a hybrid of two planning strategies:
  1. The original Graphplan-based STAN algorithm.
  2. A forward planner using a heuristic function based on the length of the relaxed plan (as in HSP and FF).
- Uses automatic of domain analysis to select between these strategies.
- The domain analysis techniques include type, and invariant detection, as well as the automatic identification of certain combinatorial optimization sub-problems
- Competed in the fully automated track.

The Competitors

10. BDDPlan

Hans-Peter Stoerr
Dresden University of Technology

- BDDPlan uses BDDs to support reasoning in the Fluent Calculus, an framework for reasoning about actions in first order logic.
- Model checking algorithms are used to do an implicit breadth first search.
- Competed in the fully automated track.

The Competitors

11. AltAlt

Biplav Srivastava
Terry Zimmerman
BinhMinh Do
XuanLong Nguyen
Zaiqing Nie
Ullas Nambiar
Romeo Sanchez

Arizona State University

- AltAlt uses effective and admissible heuristics extracted from the planning graph to drive the backward state space.
- Competed in the fully automatic track.

The Competitors

12. GRT

Ioannis Refanidis
Ioannis Vlahavas
Dimitris Vrakas

Aristotle University

- GRT (Greedy Regression Table) planner is a heuristic state space planner. Like HSP it employs a best-first search using estimated distances to the goal
- Competed in the fully automatic track.

The Competitors

13. PbR

Jose Luis Ambite
Craig Knoblock
Steve Minton

University of Southern California/Information Sciences Institute

- Planning by Rewriting (PbR) generates plans by using a set of plan rewriting rules and local search techniques to transform an easy-to-generate poor quality initial plans into a higher-quality plans.
- The rewriting rules were developed semi-automatically: some proposed by a learning algorithm, some defined manually.
- The initial plan generators were hand-coded for each domain.
- Competed in the hand tailored track.

The Competitors

14. SHOP

Dana Nau
Yue (Jason) Cao
Hector Munoz-Avila
Amnon Lotem

University of Maryland

- SHOP (Simple Hierarchical Ordered Planner) is an HTN planning system that plans for tasks in the same order that they will later be executed.
- This simplifies goal-interactions and provides a complete world-state at each step of the planning process as with forward chaining planners.
- The complete state information allows SHOP to encode effective domain specific planning algorithms and knowledge.
- Competed in the hand tailored track.

The Competitors

14. TALPlanner

Jonas Kvarnstrom
Patrick Doherty
Patrik Haslum

Linkoping University

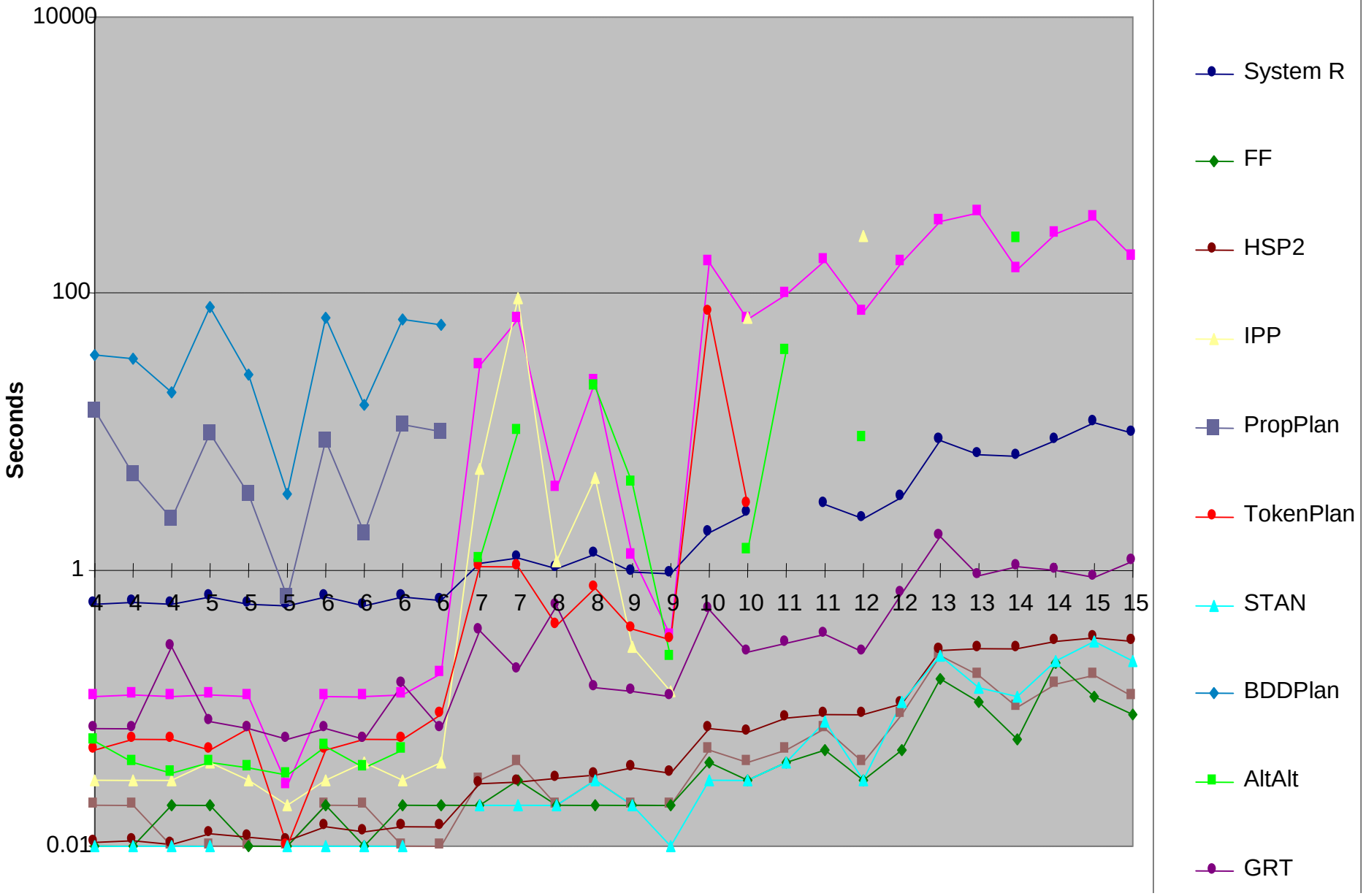
- TALplanner is a forward-chaining planner based on the TLPlan system (Bacchus & Kabanza).
- Domain-dependent search control knowledge expressed declaratively as formulas of a temporal logic and used to control forward chaining.
- TALplanner uses TAL a narrative temporal logic for reasoning about action and change.
- Competed in the hand tailored track.

The Results

Domain 1: Logistics World

- Move a set of packages between locations using trucks within the same city and airplanes between cities.
- Limited interaction between goals.

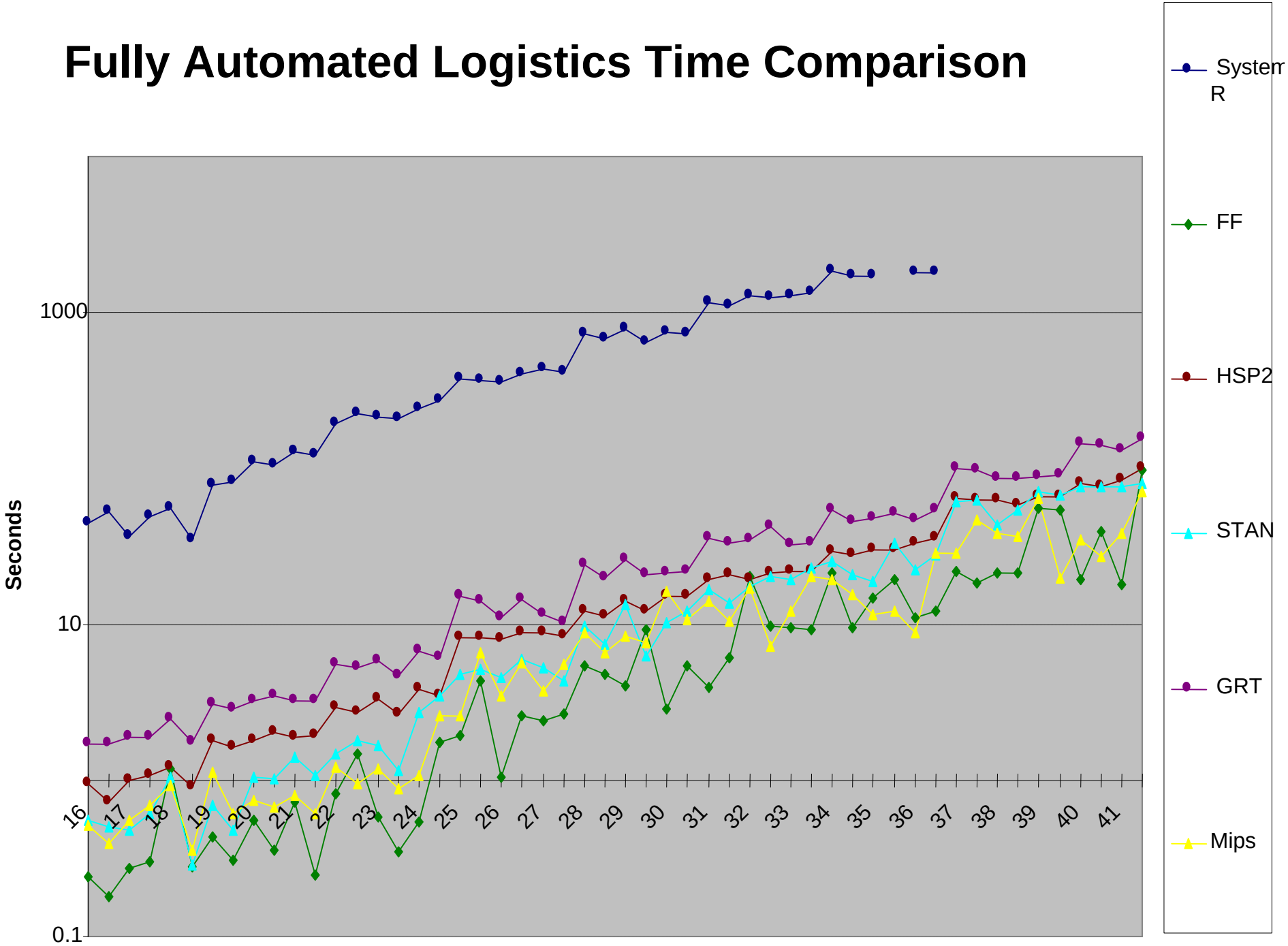
Fully Automated Logistics Time Comparison



Planners doing well enough to scale to bigger problems:

- System R
- GRT
- HSP2
- Stan
- Mips
- FF

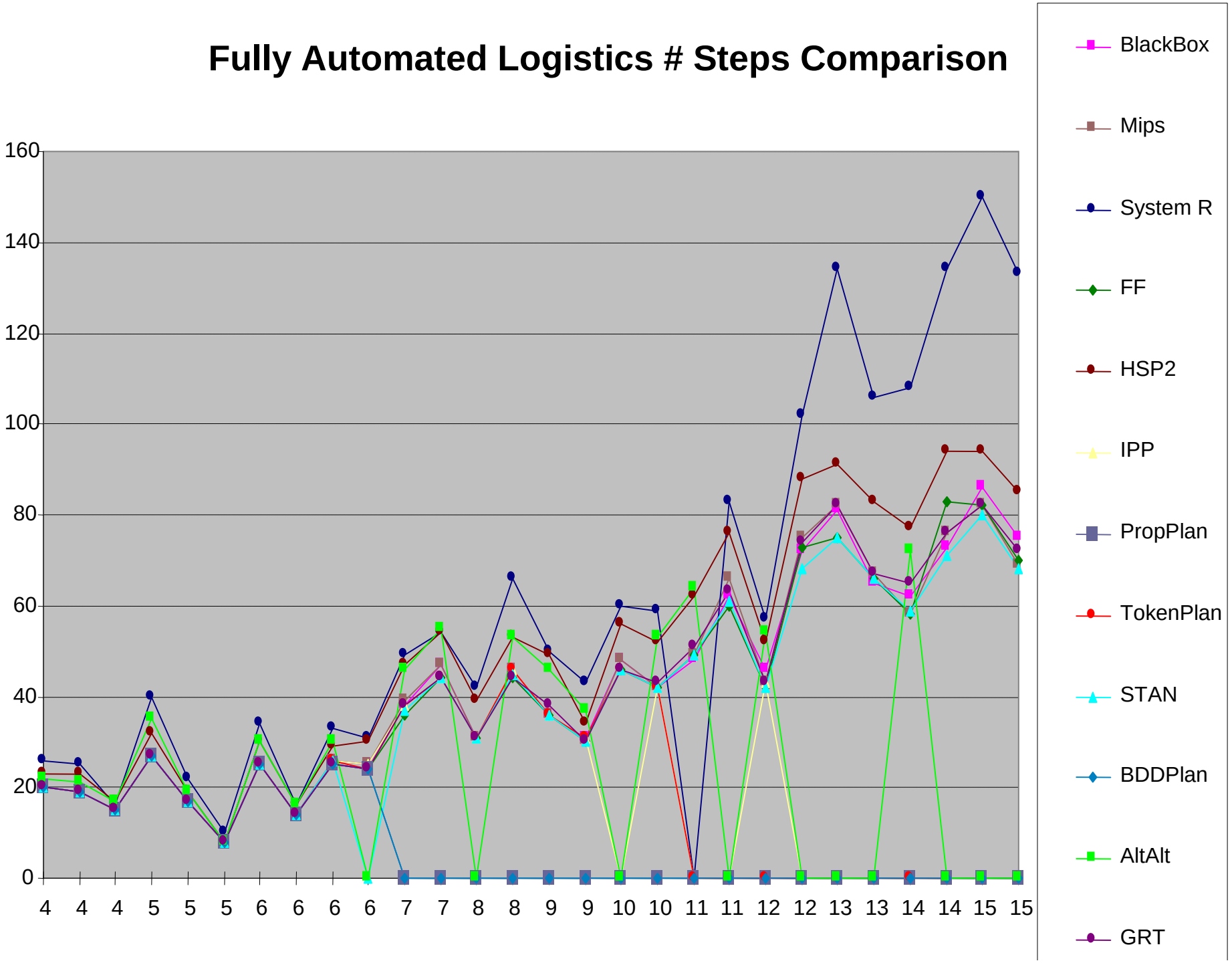
Fully Automated Logistics Time Comparison



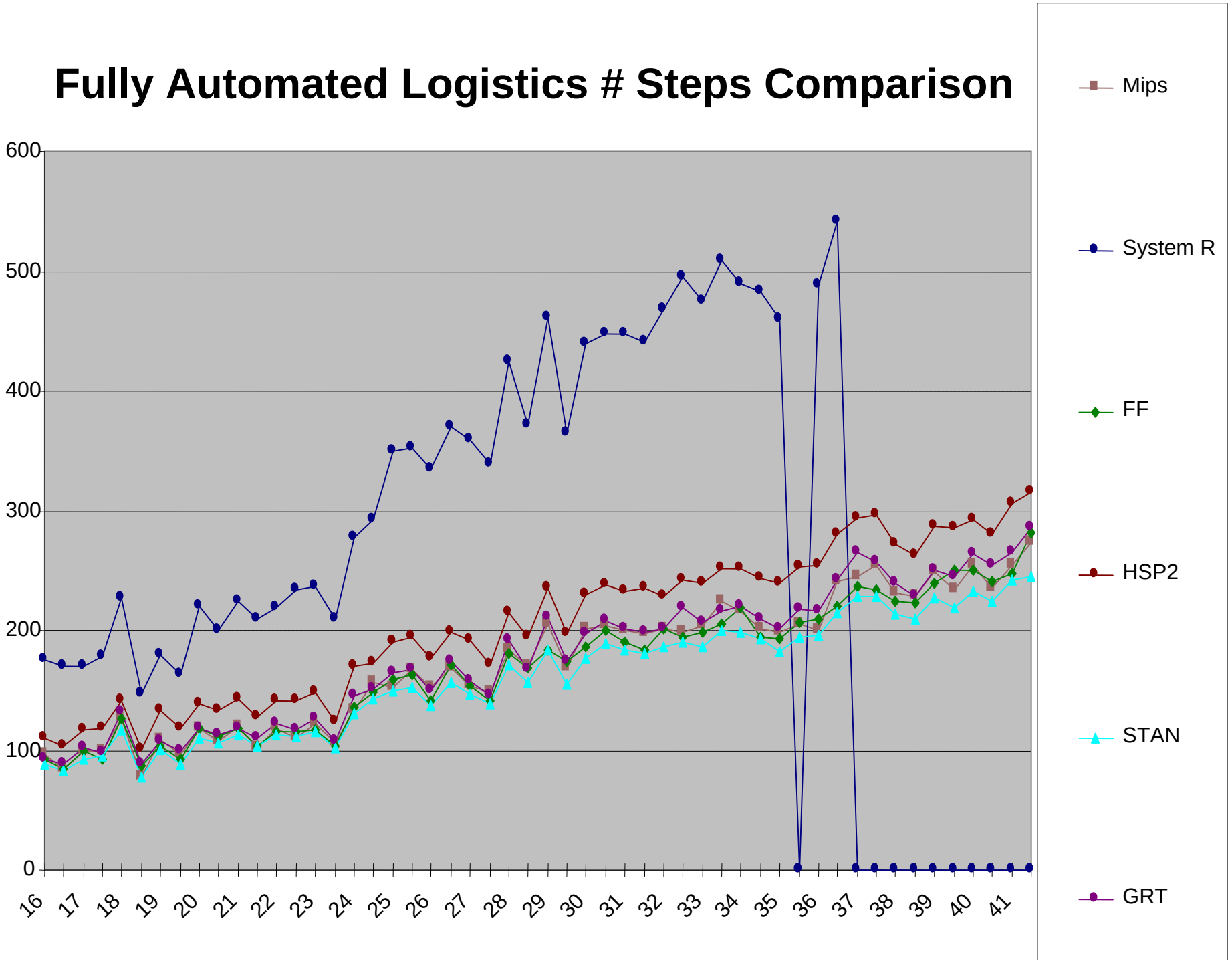
Logistics Domain Time

- FF, MIPS a bit better
- Stan and HSP2
- Then GRT

Fully Automated Logistics # Steps Comparison



Fully Automated Logistics # Steps Comparison

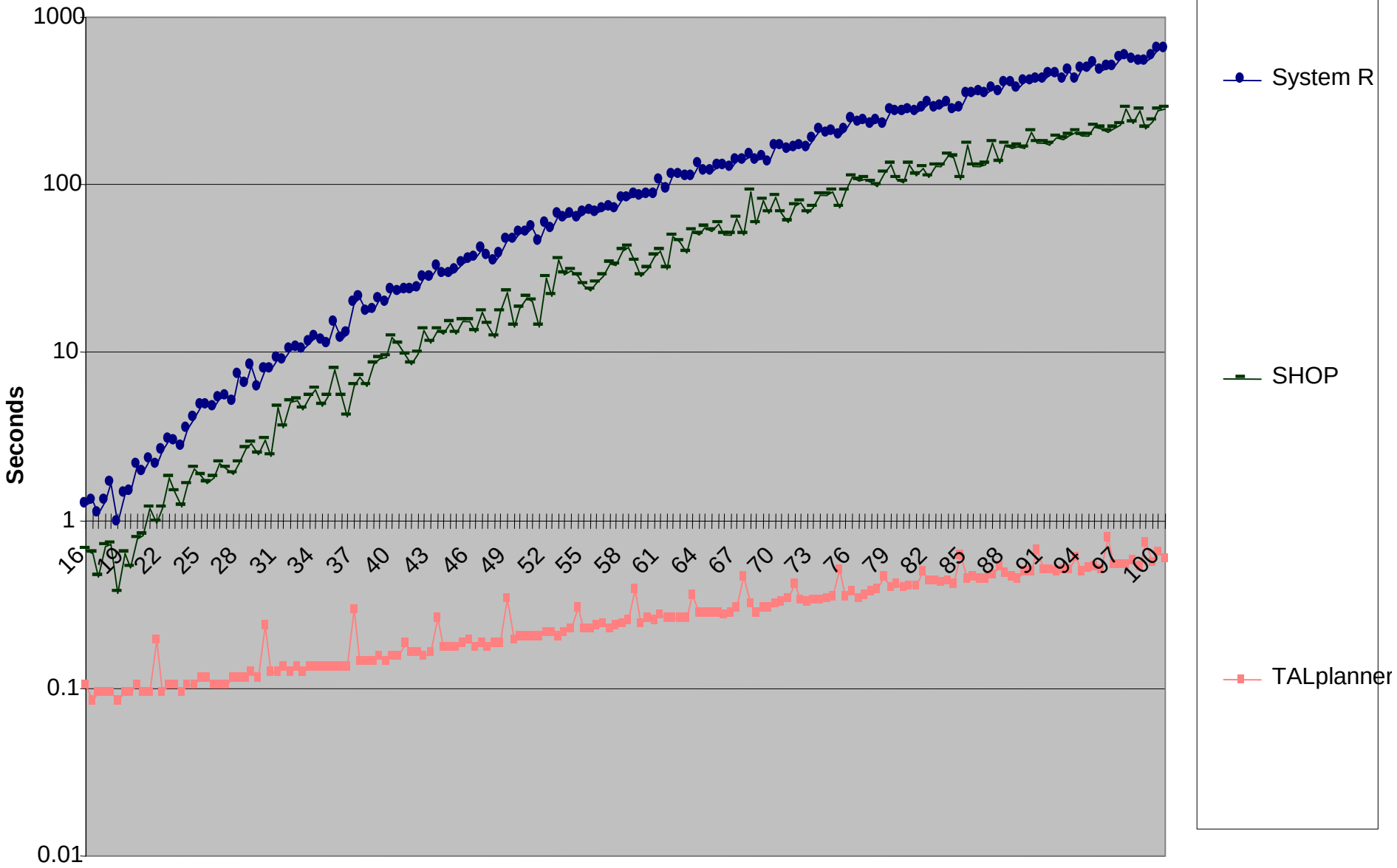


Logistics Plan Length

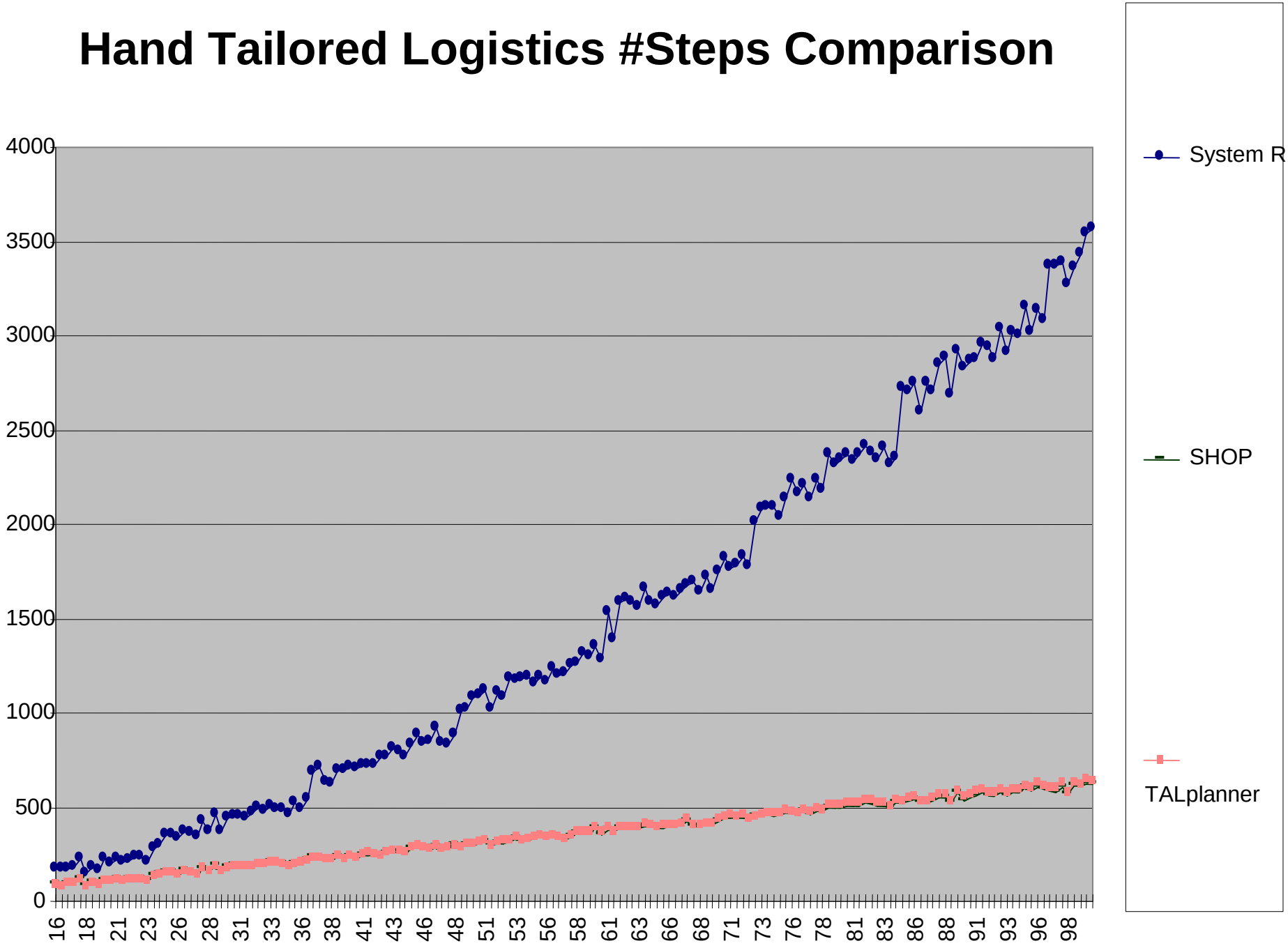
- Stan is generating the shortest plans
- GRT, MIPS, FF about the same
- HSP a bit worse
- System R generates very long plans in this domain.

Hand Tailored

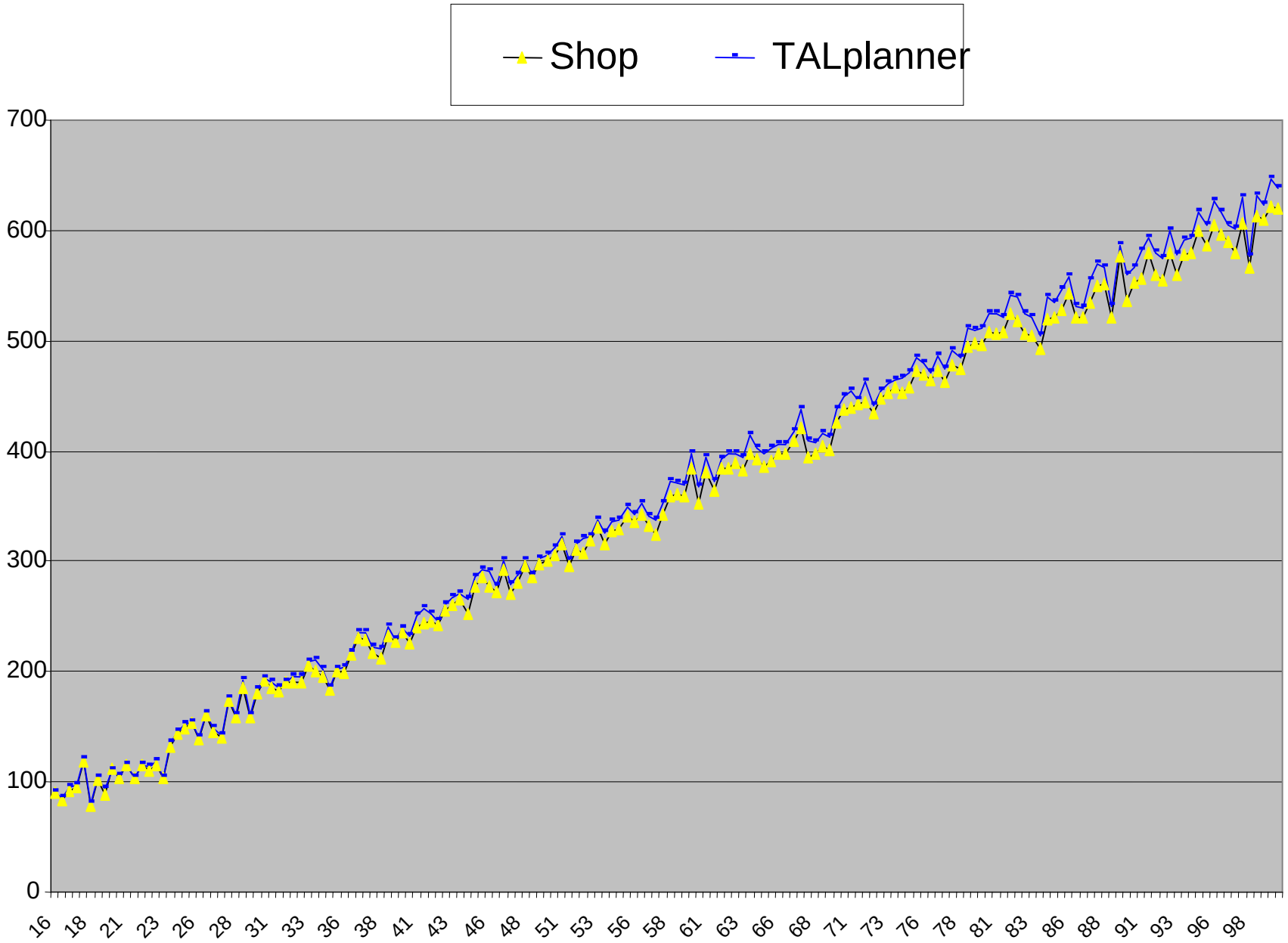
Hand Tailored Logistic Time Comparison



Hand Tailored Logistics #Steps Comparison



Hand Tailored Logistics # Steps Comparison



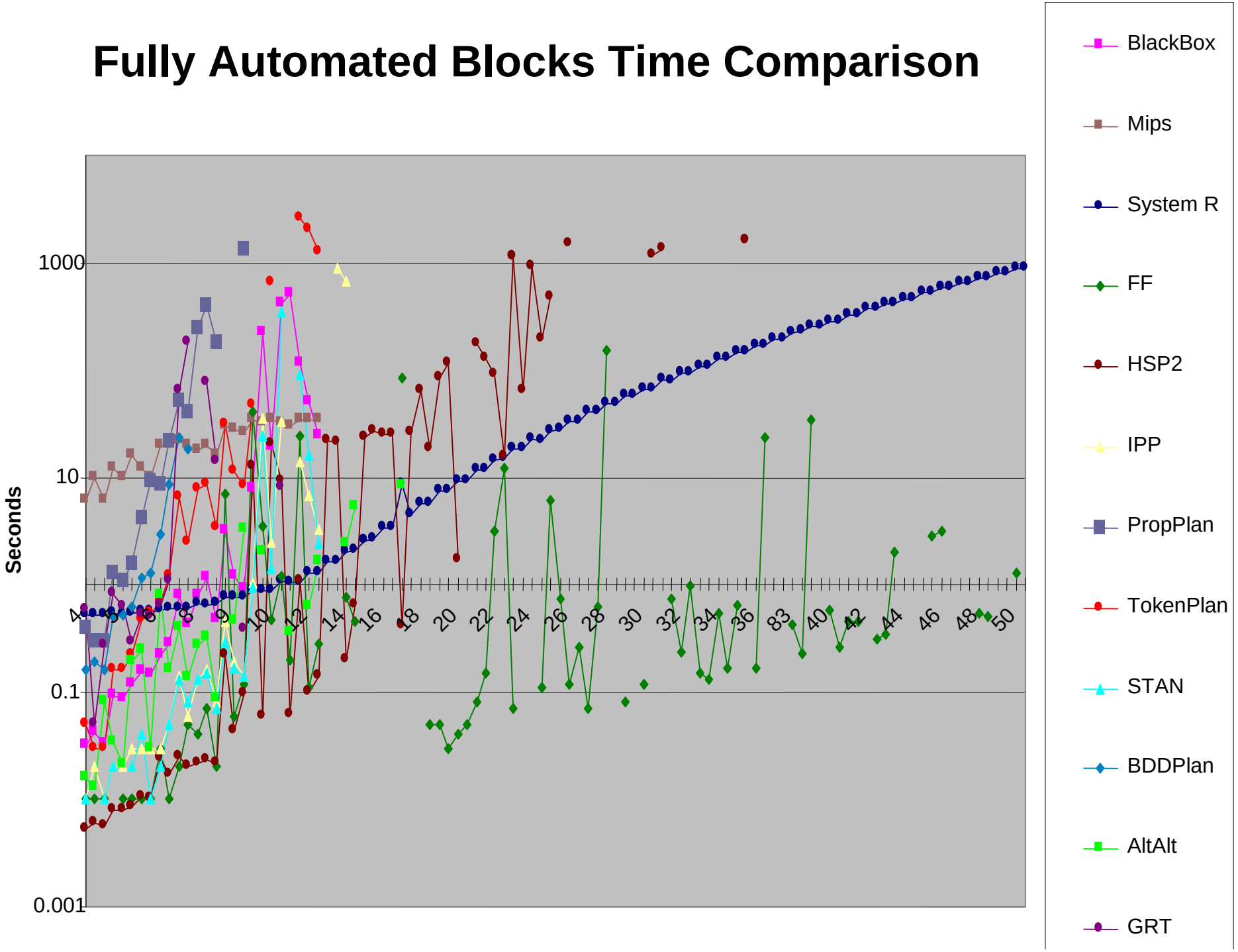
Hand Tailored Logistics

- TALplan is extremely fast (the largest problems in less than a second
- Shop and then System R
- Shop and TALplan are generating short plans (with Shop a bit better)
- System R is generating very long plans

Domain 2: Blocks World

- Stack a set of blocks.
- A high degree of interaction between goals.
- Easy for people, can be hard for automated planners.

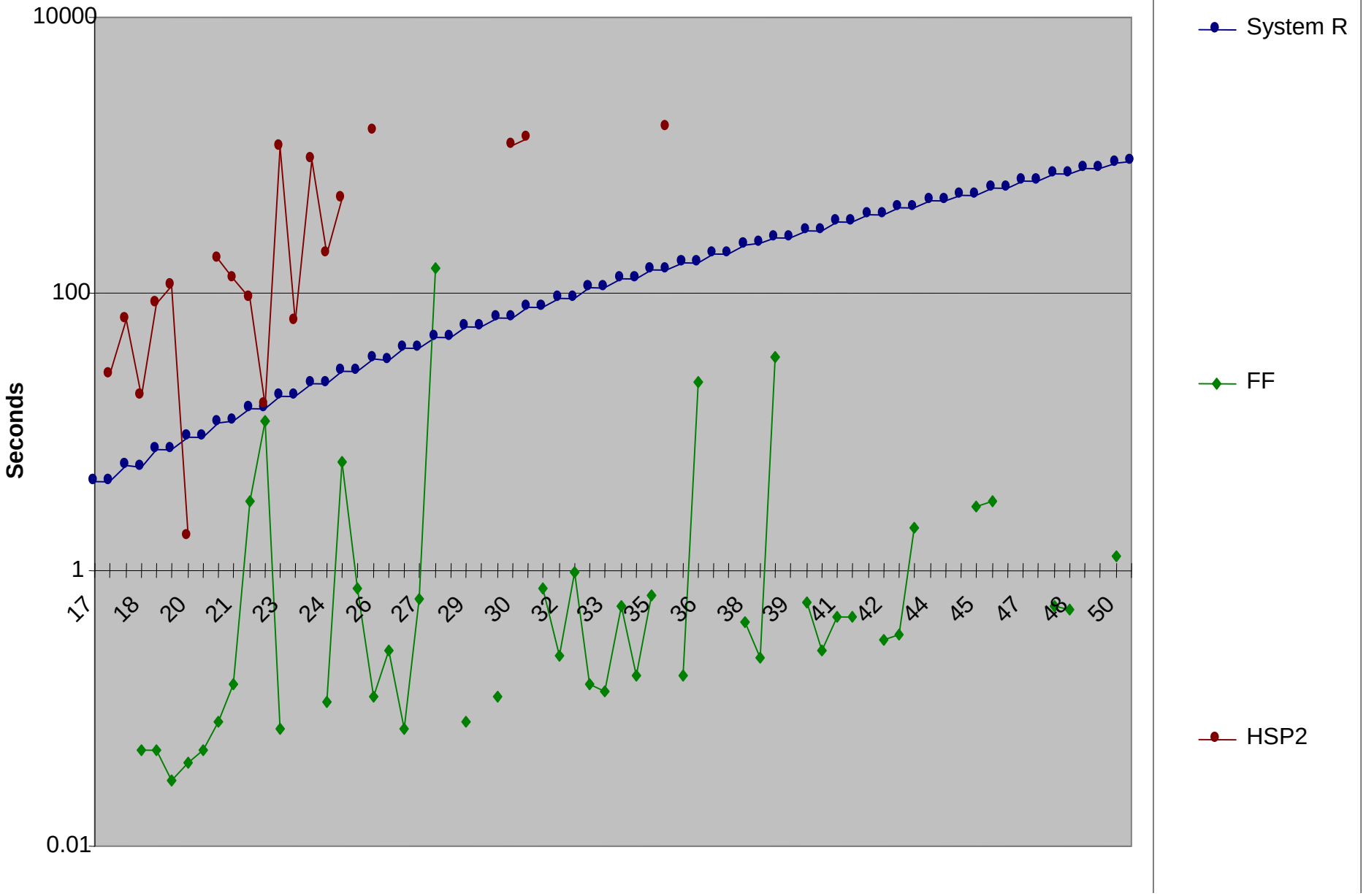
Fully Automated Blocks Time Comparison



Blocks

- A wide variation of performance.
- Only FF, System R, and HSP2 are able to solve the larger problems.

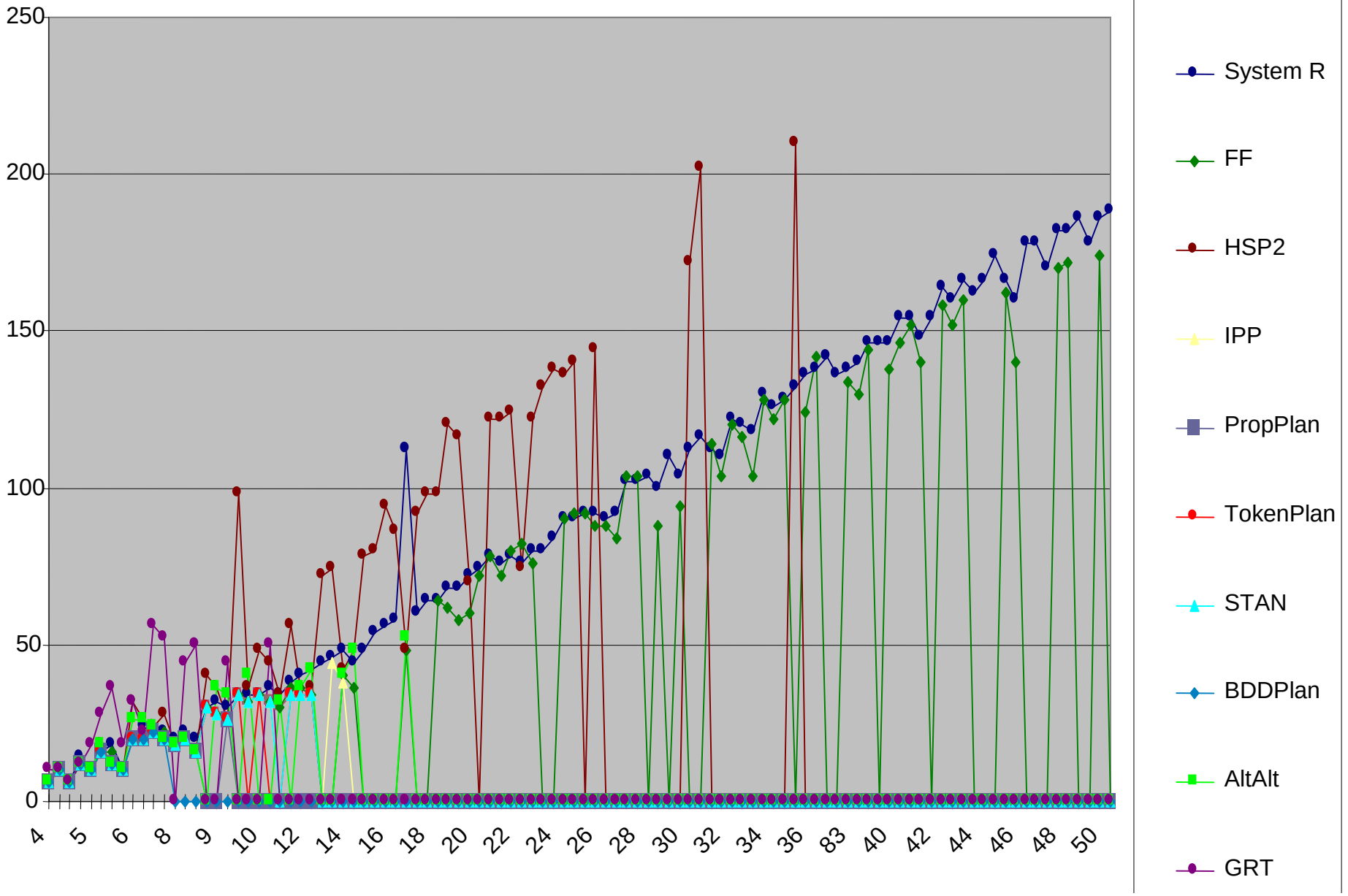
Fully Automated Blocks Time Comparison



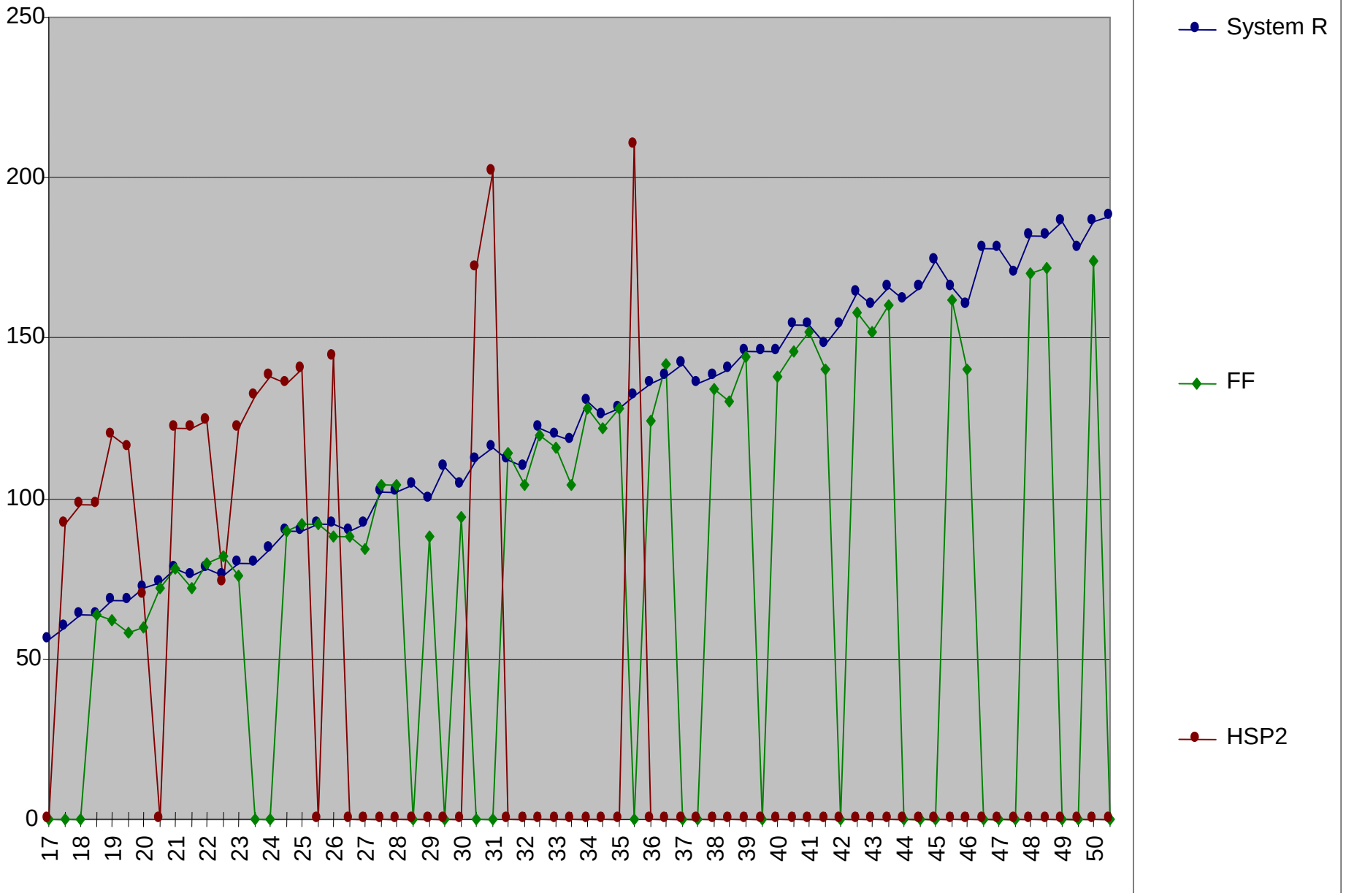
Blocks

- System R scales very consistently in this domain.
- FF can occasionally solve large problems fast, but it has many misses.

Fully Automated Blocks # Steps Comparison



Fully Automated Blocks # Steps Comparison

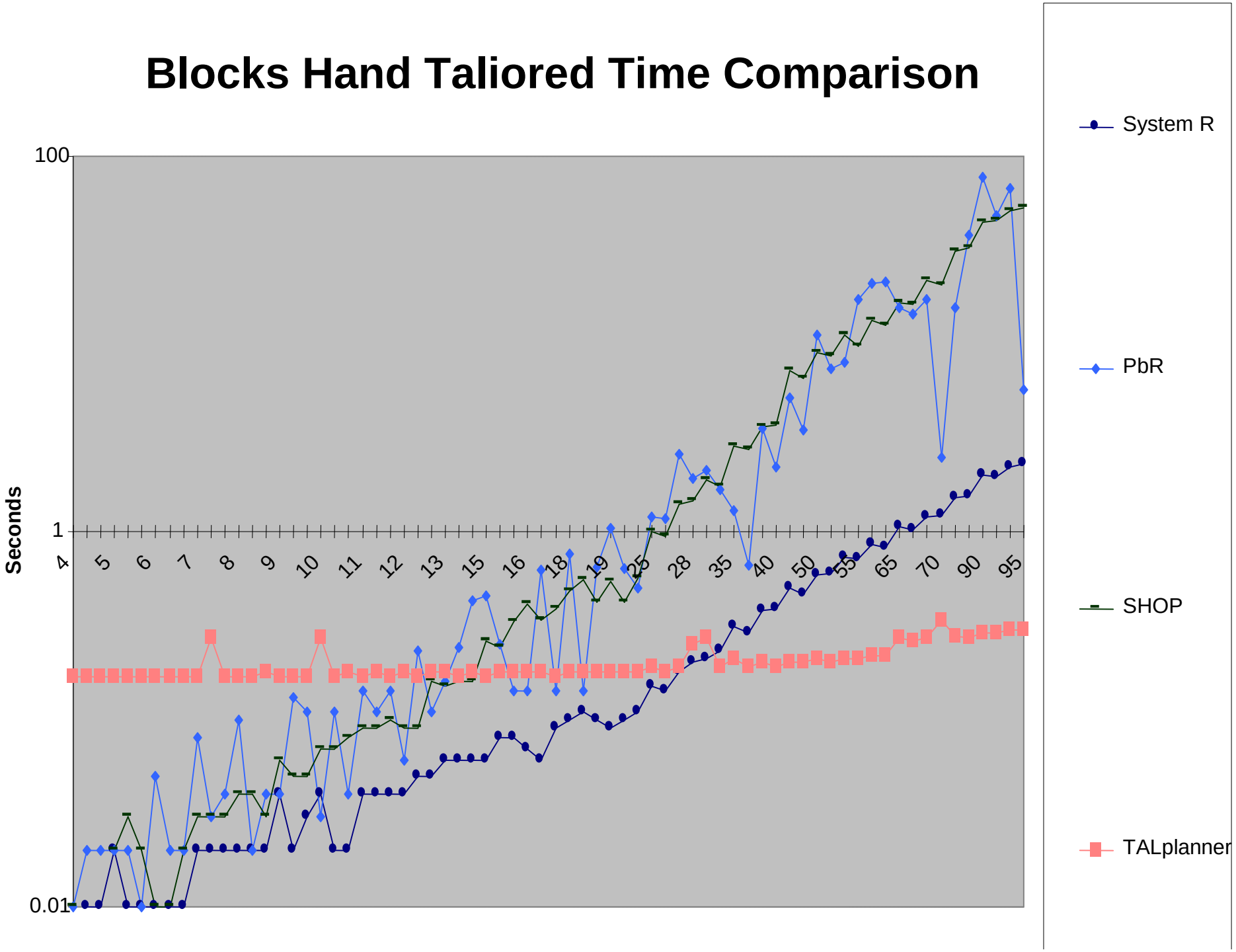


Blocks

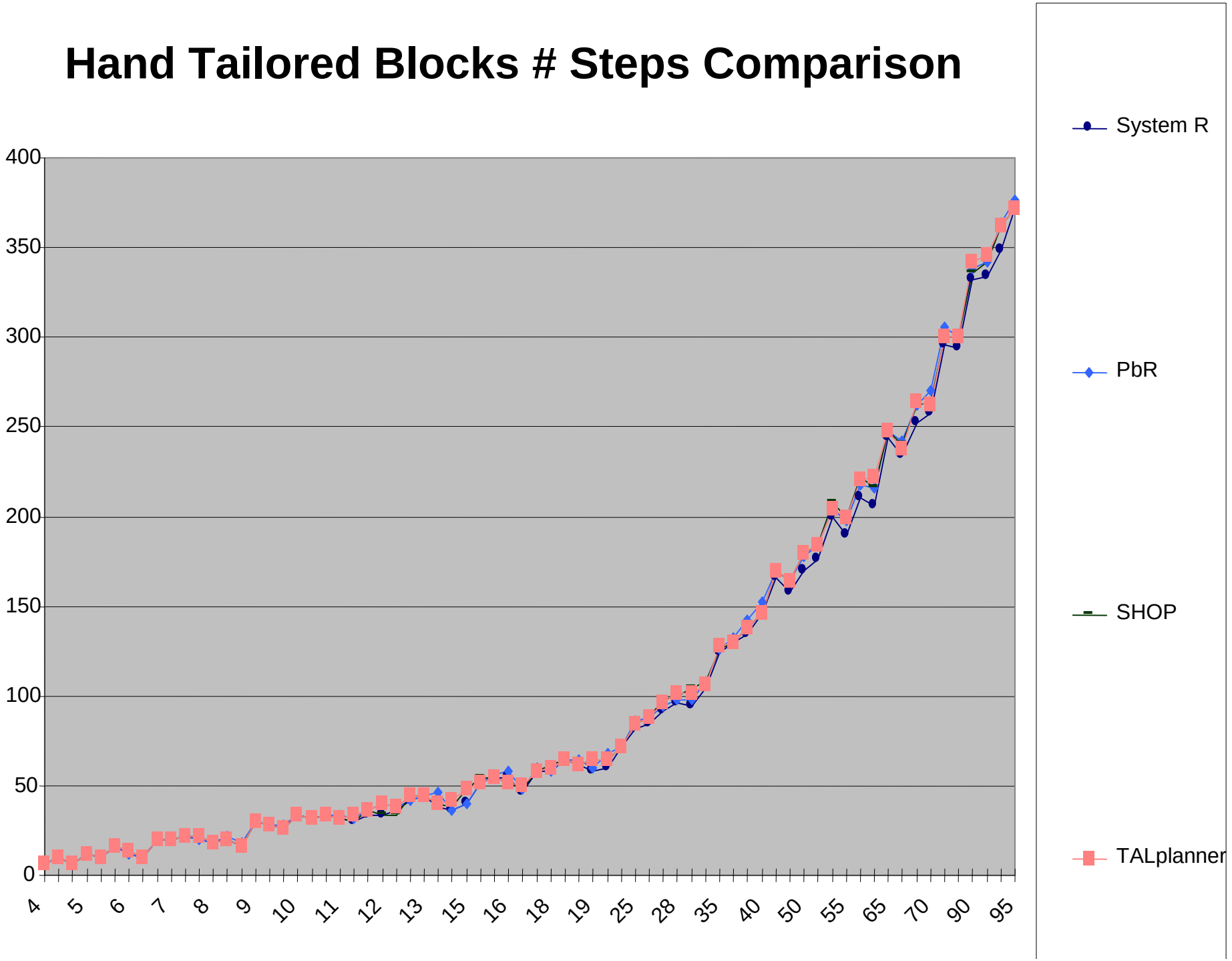
- System R generates short plans.
- FF when it succeeds also generates short plans.
- HSP2 can generate very long plans in this domain.

Blocks Hand Tailored

Blocks Hand Taliored Time Comparison



Hand Tailored Blocks # Steps Comparison

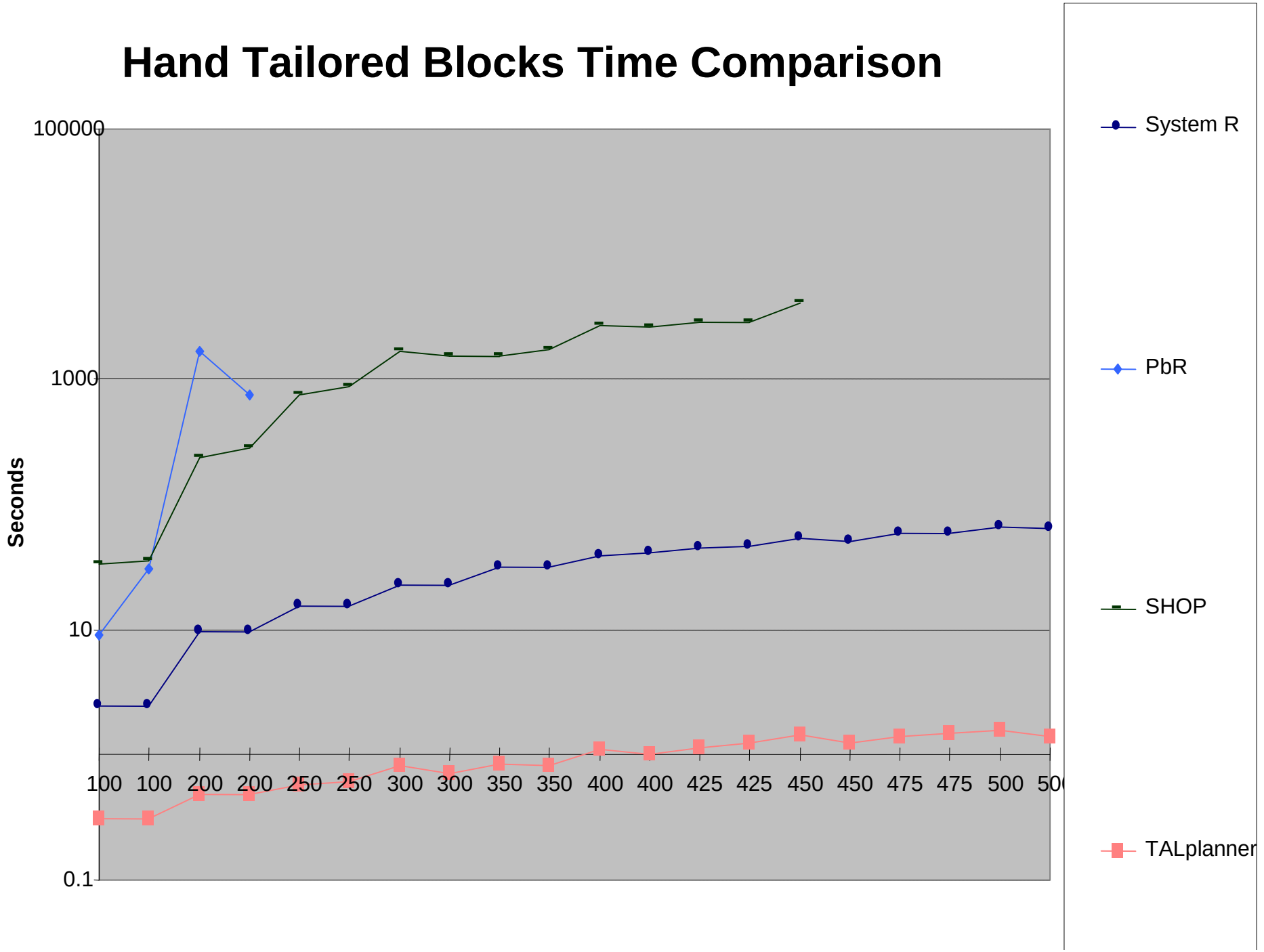


Blocks—Hand Tailored

- TALPlan is very fast.
- System R, then Shop and PBR.
- They all generate plans of a similar length.

Blocks Some Harder Problems

Hand Tailored Blocks Time Comparison



Blocks

- TALplan scales very well in this domain solving 500 block problems in about 1.5 seconds.
- System R also scales well enough to solve the hardest problems of the .

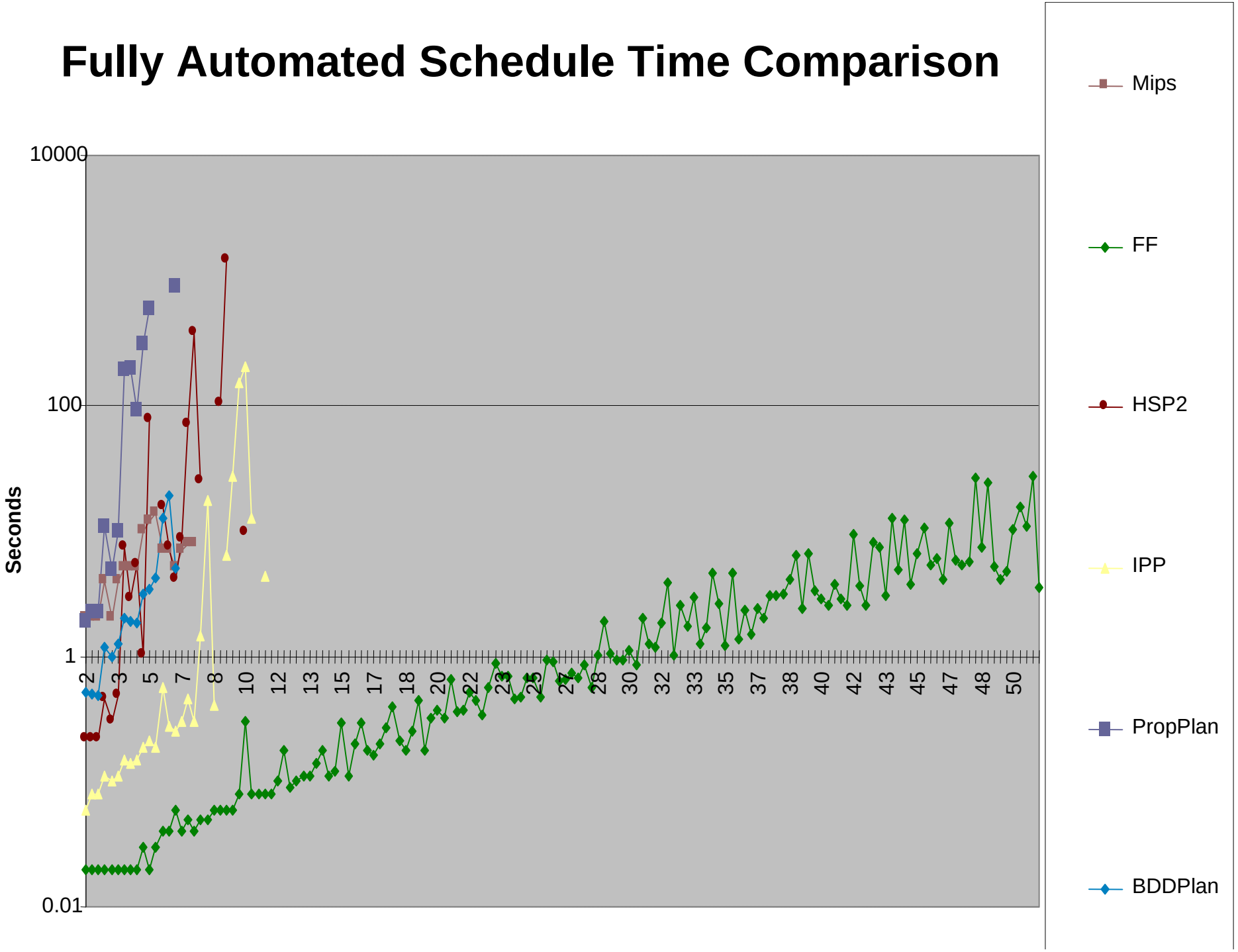
Domain 3: Schedule World

- Machine a collection of parts.
- Goals are mostly non-interacting, but they compete for “resources” (time on machines) and on the same part different goals clobber other goals.
- Originally a Prodigy domain.

Schedule World

- This is an ADL domain with many actions requiring conditional effects.
- Only Mips, FF, HSP2, IPP, PropPlan, and BDDPlan could deal with this domain.

Fully Automated Schedule Time Comparison

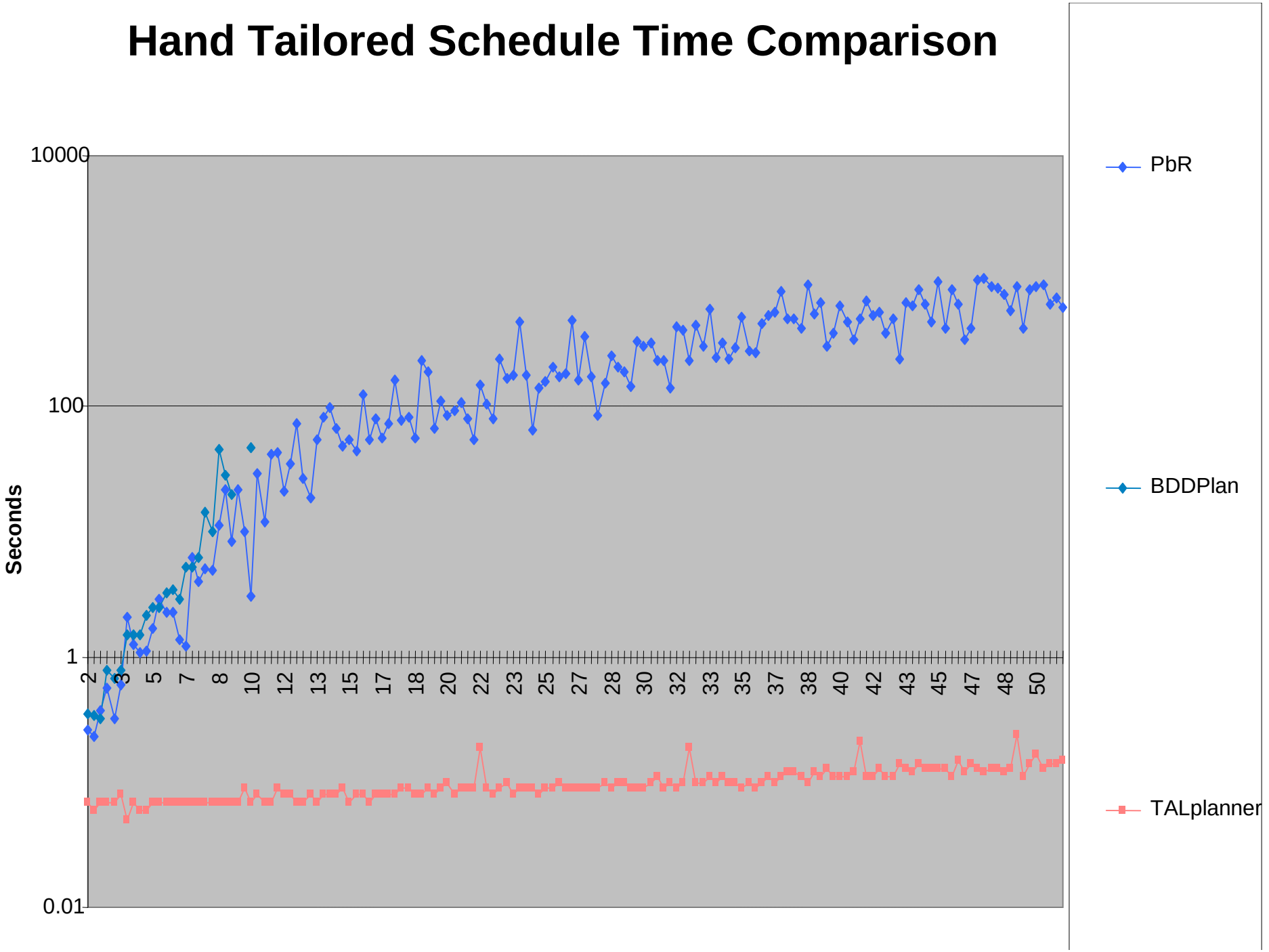


Schedule

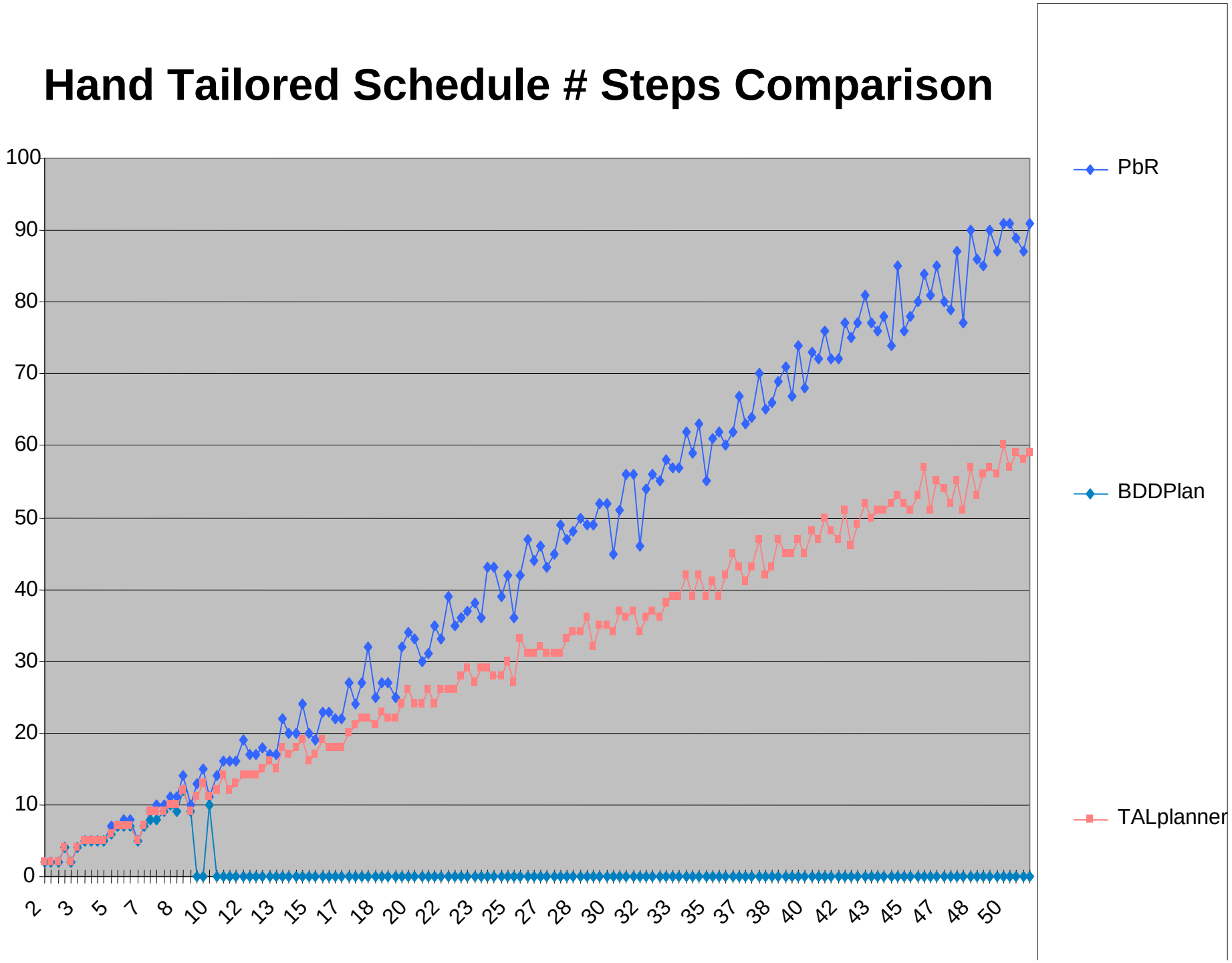
- FF is the only planner that scales to the harder problems on this domain.
- The length of the solutions are roughly comparable

Schedule—Hand Tailored

Hand Tailored Schedule Time Comparison



Hand Tailored Schedule # Steps Comparison



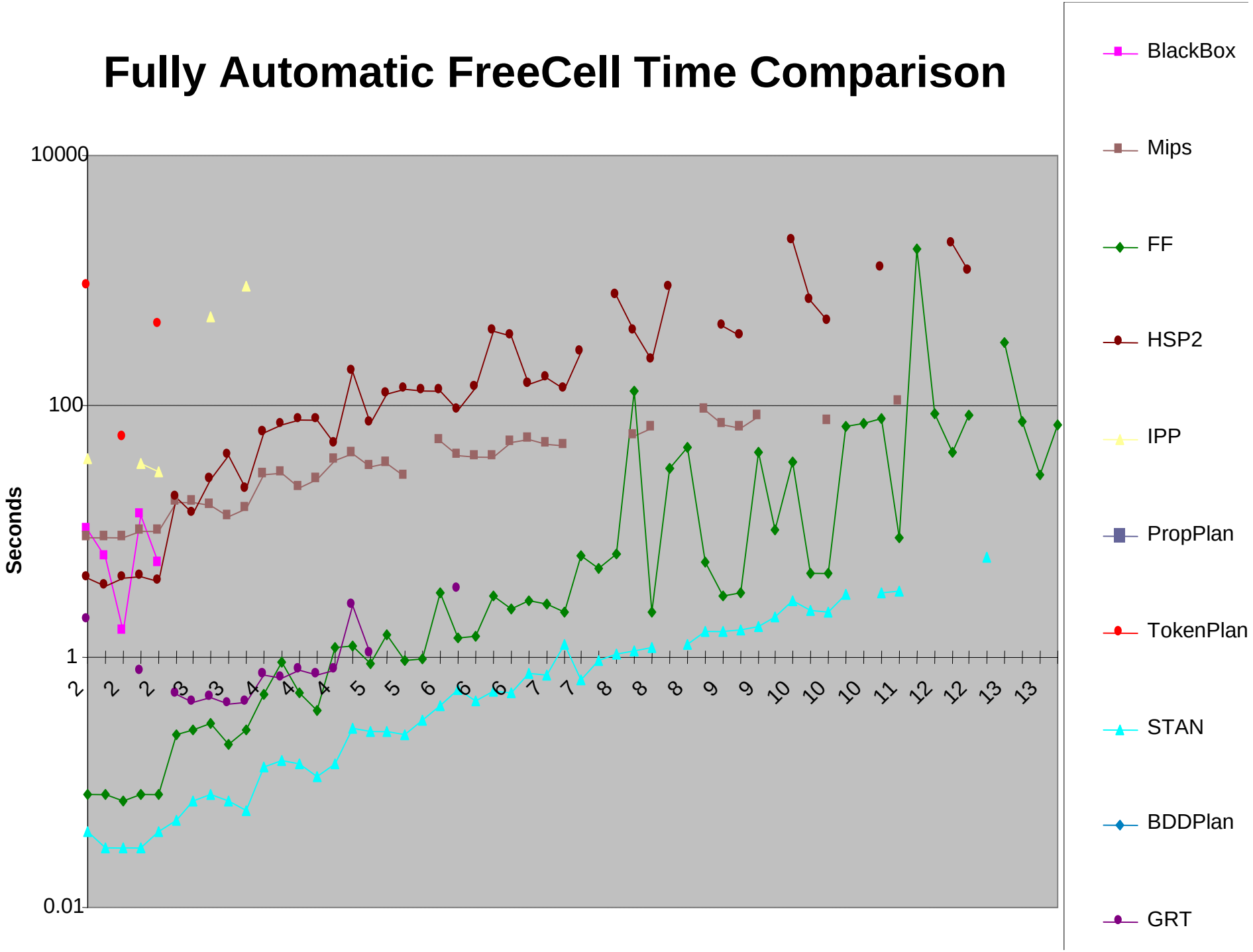
Schedule—Hand Tailored

- TALplan is generating short solutions in about 0.15 seconds on the largest problems.
- PbR takes longer and generates inferior plans.
- Interestingly FF is taking 10-25 seconds on the largest problems, generating slightly longer solutions than TALplan, but fully automatically.

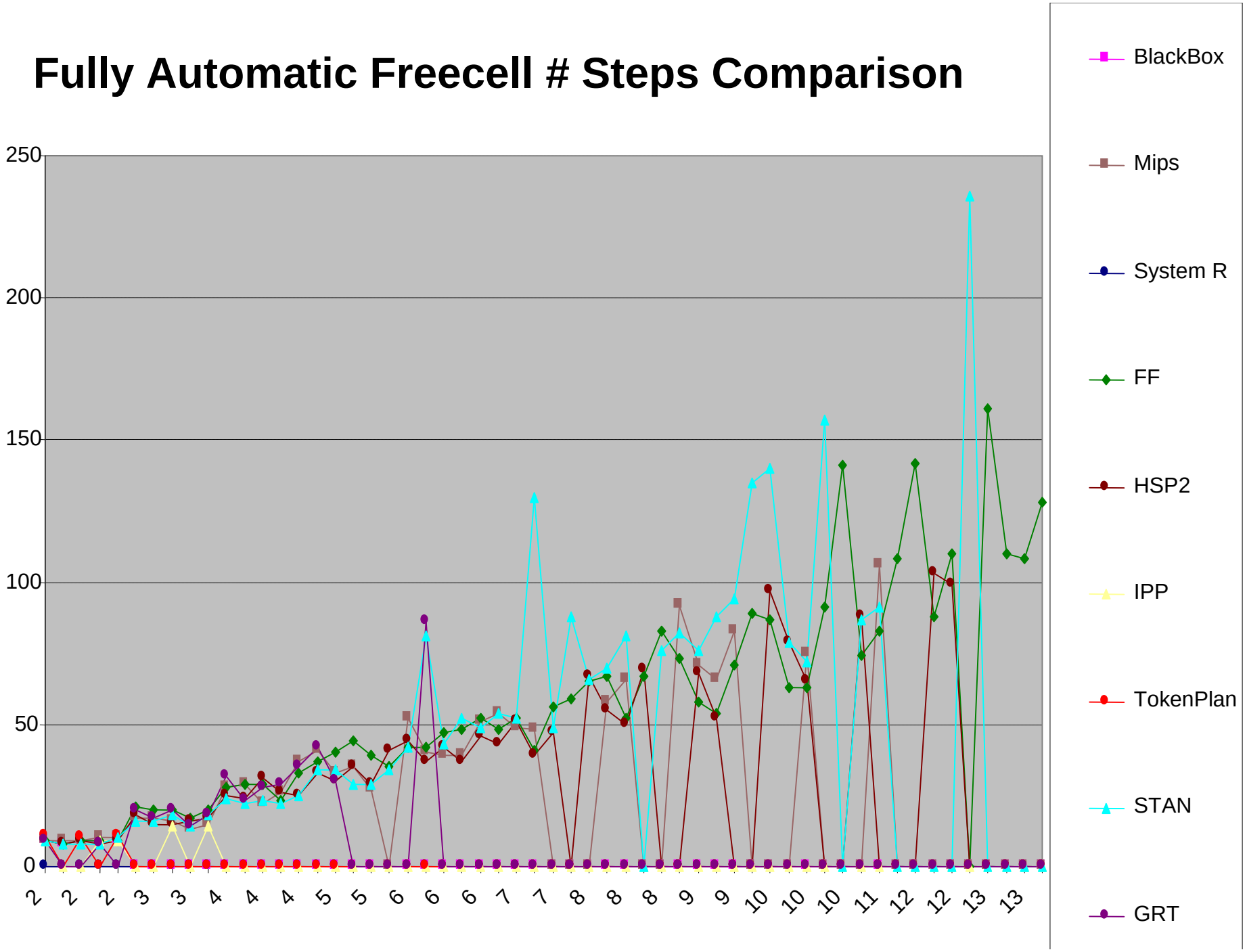
Domain 4: Freecell World

- Freecell is a solitaire card game that comes with Microsoft Windows.
- Freecell demo.

Fully Automatic FreeCell Time Comparison



Fully Automatic Freecell # Steps Comparison

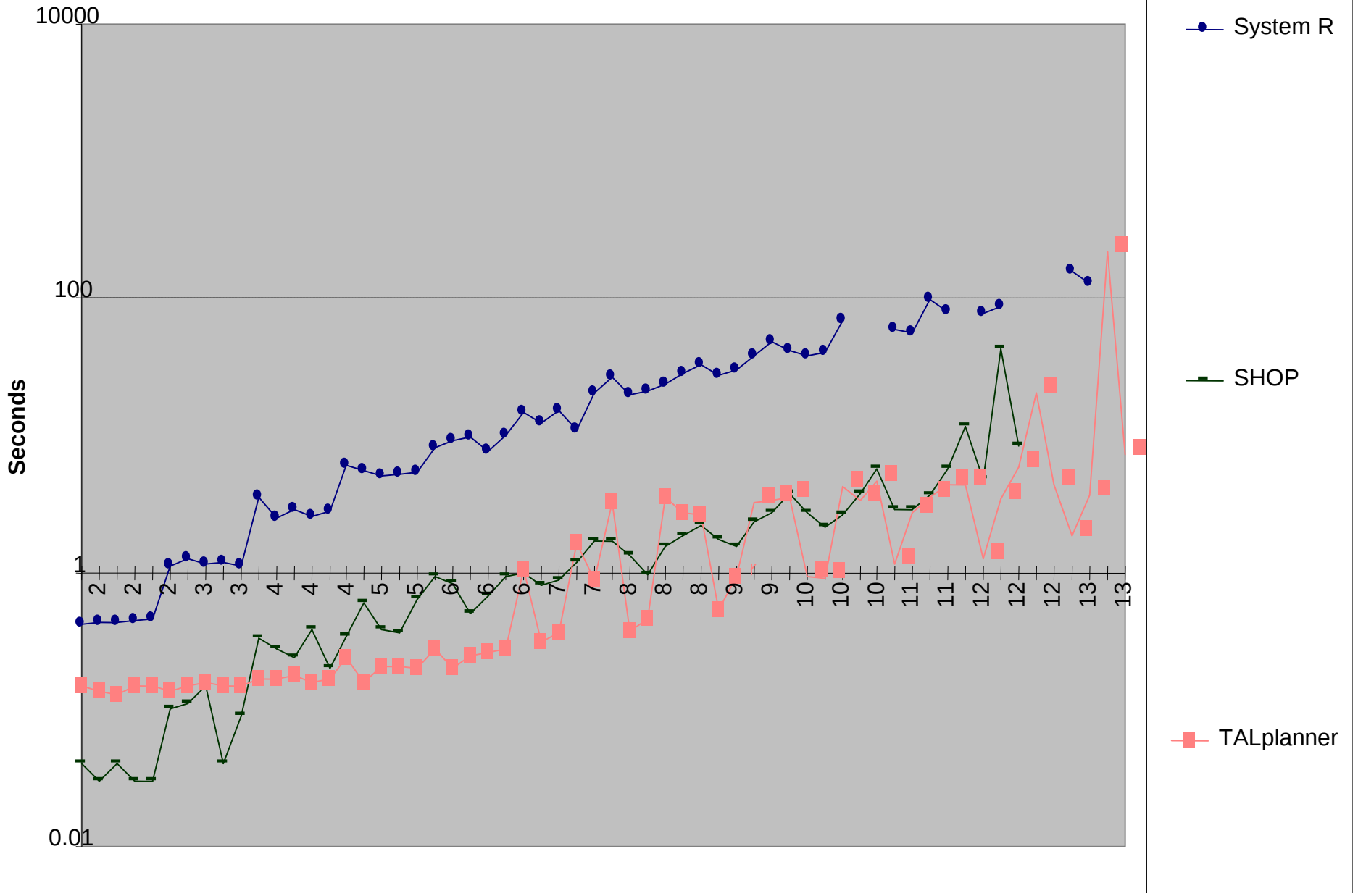


Freecell

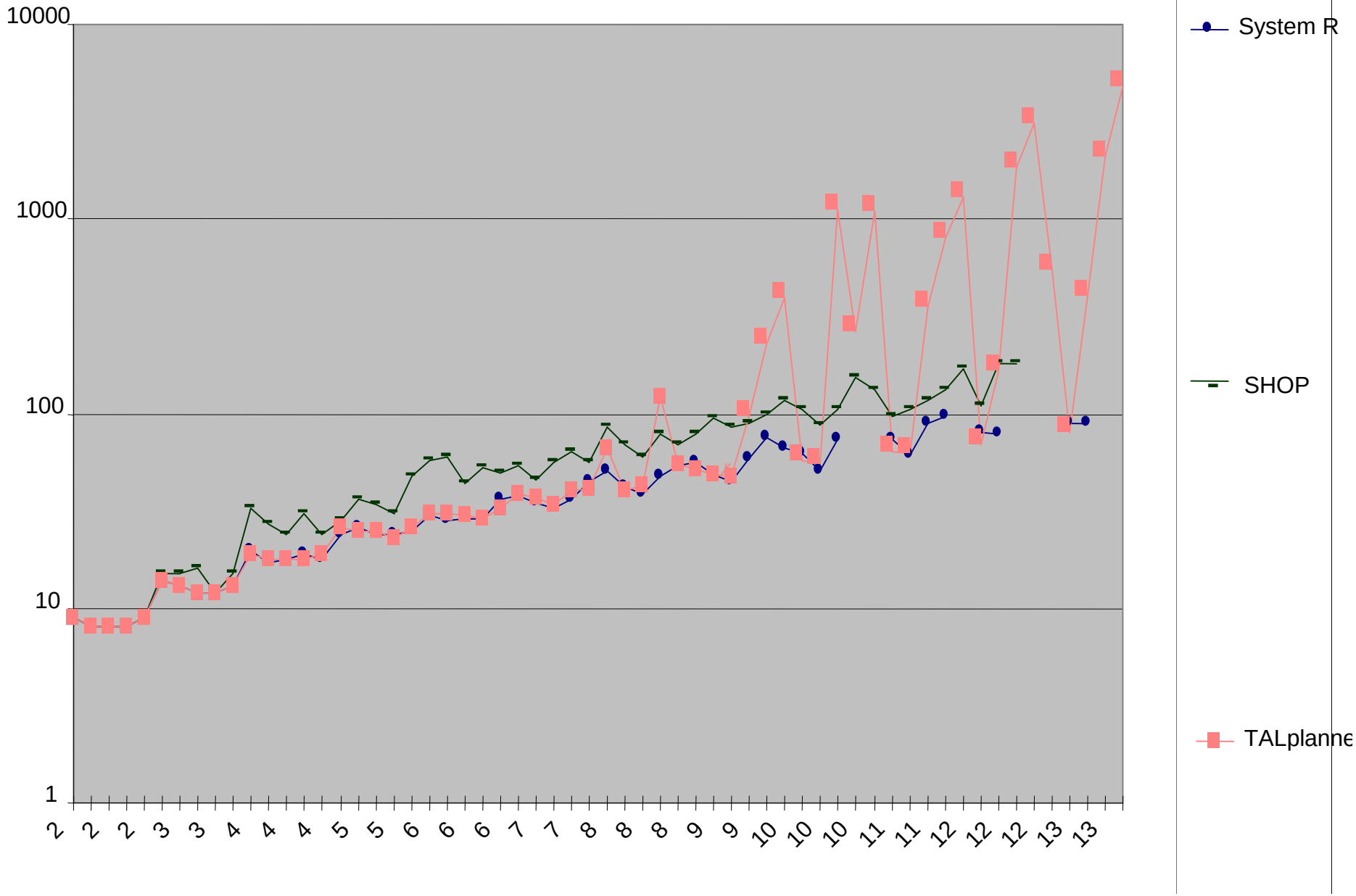
- Stan plan takes the least time but does not solve the hardest problems, and generates long solutions.
- FF is best at the harder problems, but cannot solve all the problems
- HSP2 also solves some larger problems (but takes a long time on them)

Freecell—Fully Automatic

Hand Tailored Freecell Time Comparison



Hand Tailored Freecell # Steps Comparison LOG SCALE



Freecell

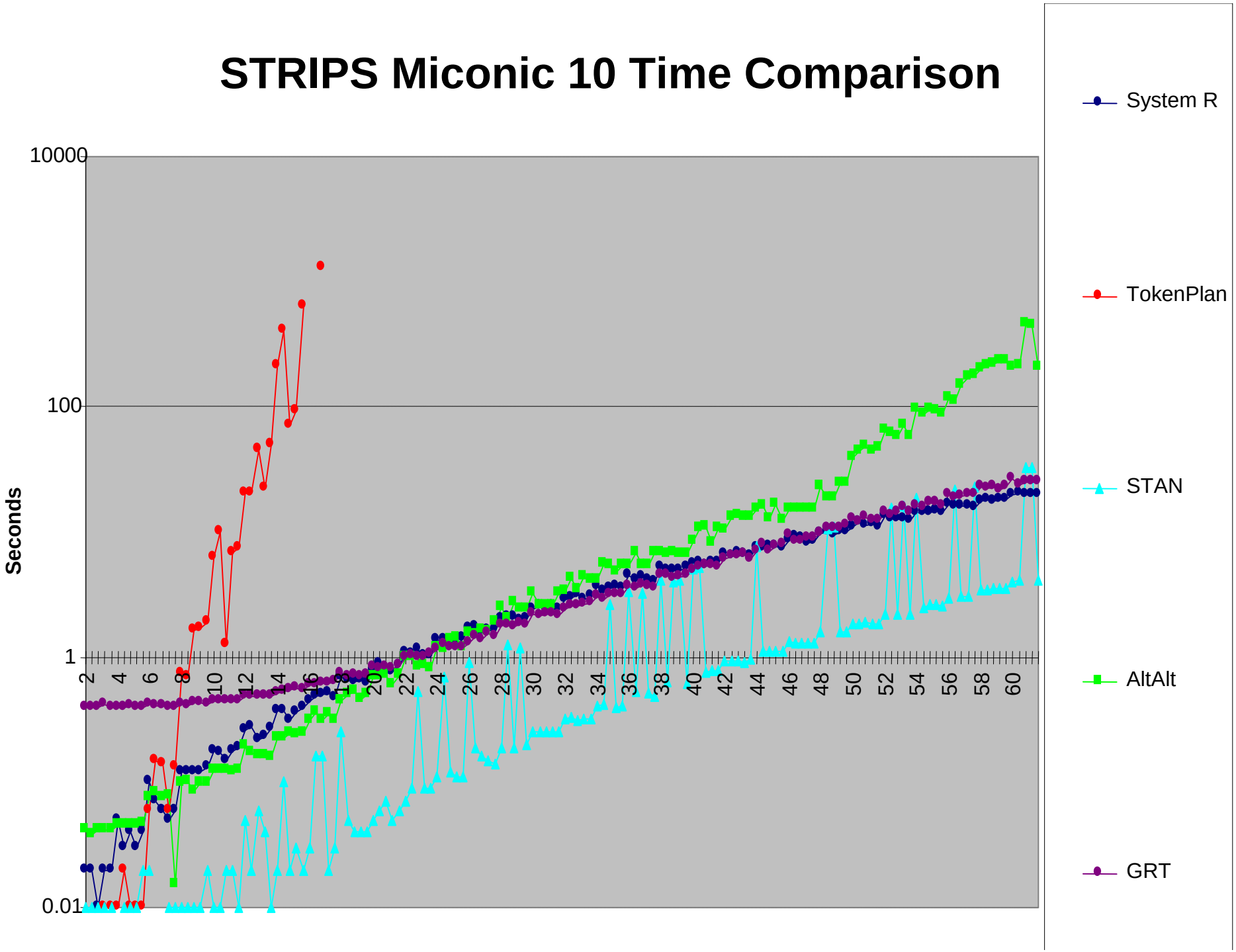
- Talplan can be fast and solves all of the problems, but it can generate very long plans 4500 steps.
- Shop generates reasonable plans in a reasonable amount of time. System R takes longer.
- None of these planners perform that much better than the fully automatic planners in this domain.

Domain 5: Mic10 Elevator World

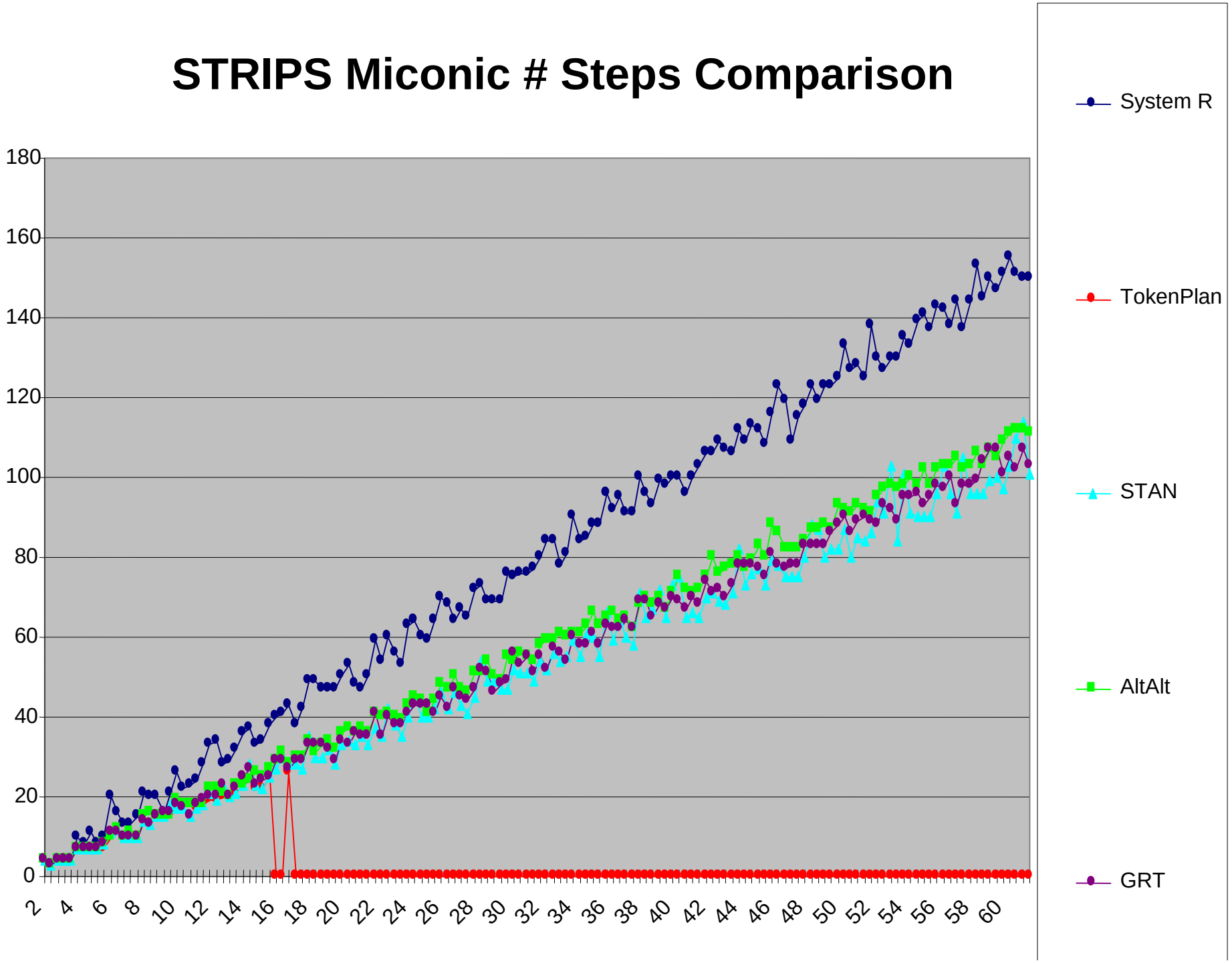
- Based on the Miconic-10 Elevator controller developed by Schindler Lifts Ltd.
- Contributed by Jana Koehler.
- Problems involve controlling a sophisticated elevator to move passengers to their destination.
- Various constraints on movement, including priority passengers, passengers that must go non-stop, passengers that must be accompanied.

Miconic 10

- Four Separate problems:
- A simple strips version.
- A simplified ADL version.
- Jana's original ADL version.
- The original ADL version plus the constraint that only 6 people can be on board the elevator at a time.


STRIPS Miconic 10 Time Comparison
Seconds
10000
100
1
0.01
2
4
6
8
10
12
14
16
18
20
22
24
26
28
30
32
34
36
38
40
42
44
46
48
50
52
54
56
58
60
System R
TokenPlan
STAN
AltAlt
GRT

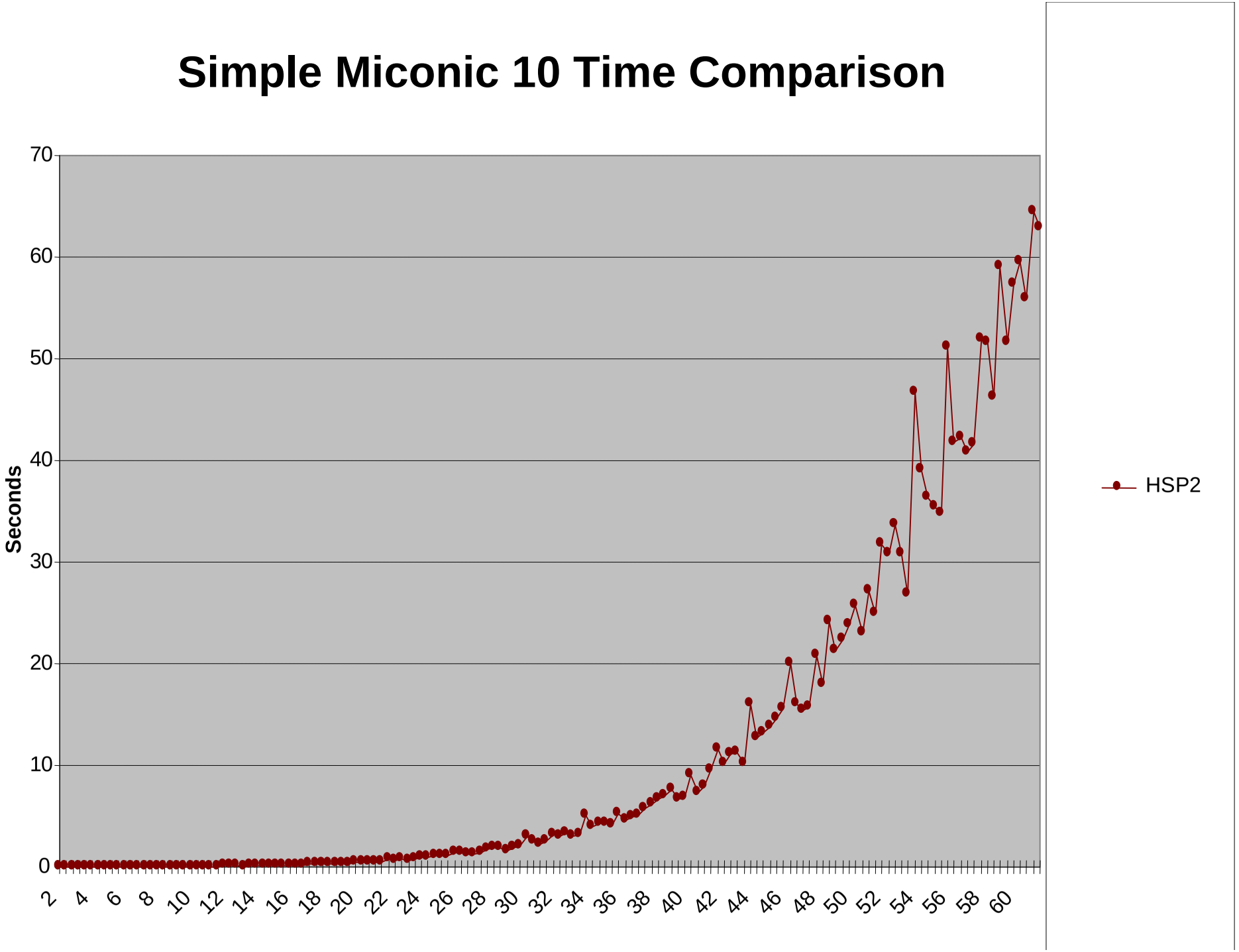
STRIPS Miconic # Steps Comparison



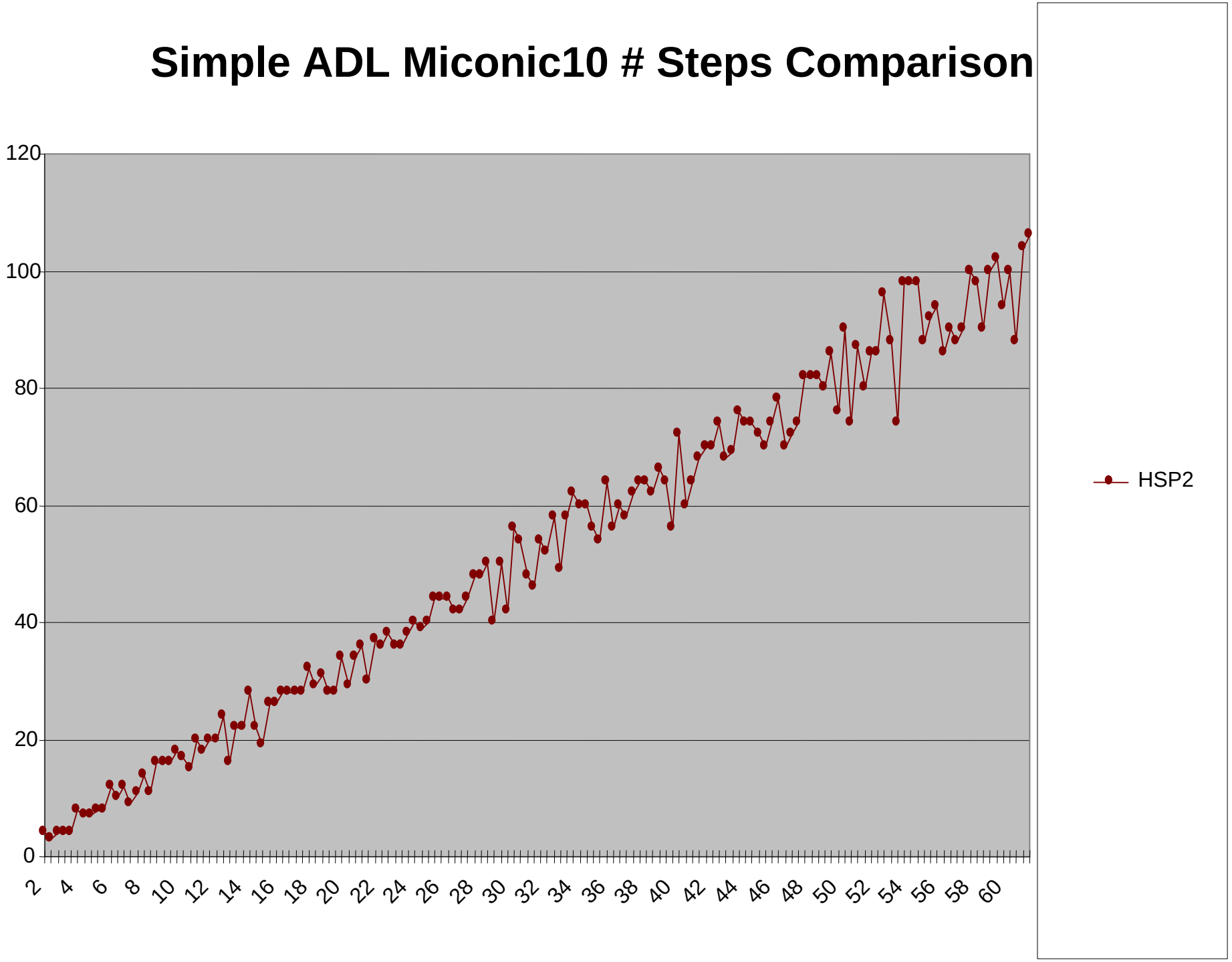
Strips Miconic10

- STAN is slightly faster and generates slightly shorter solutions.
- GRT also does well on both criteria.

Simple Miconic 10 Time Comparison



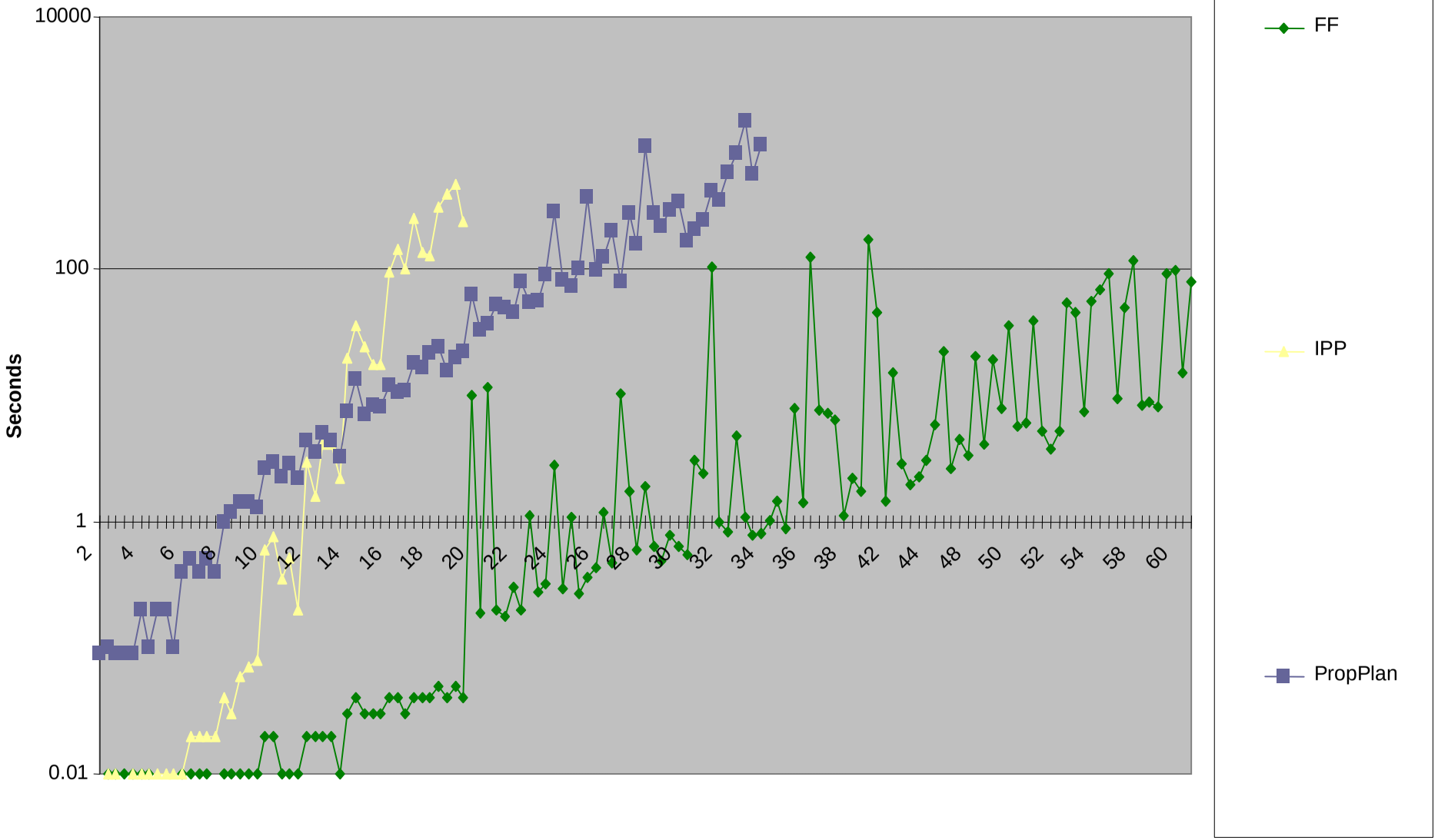
Simple ADL Miconic10 # Steps Comparison



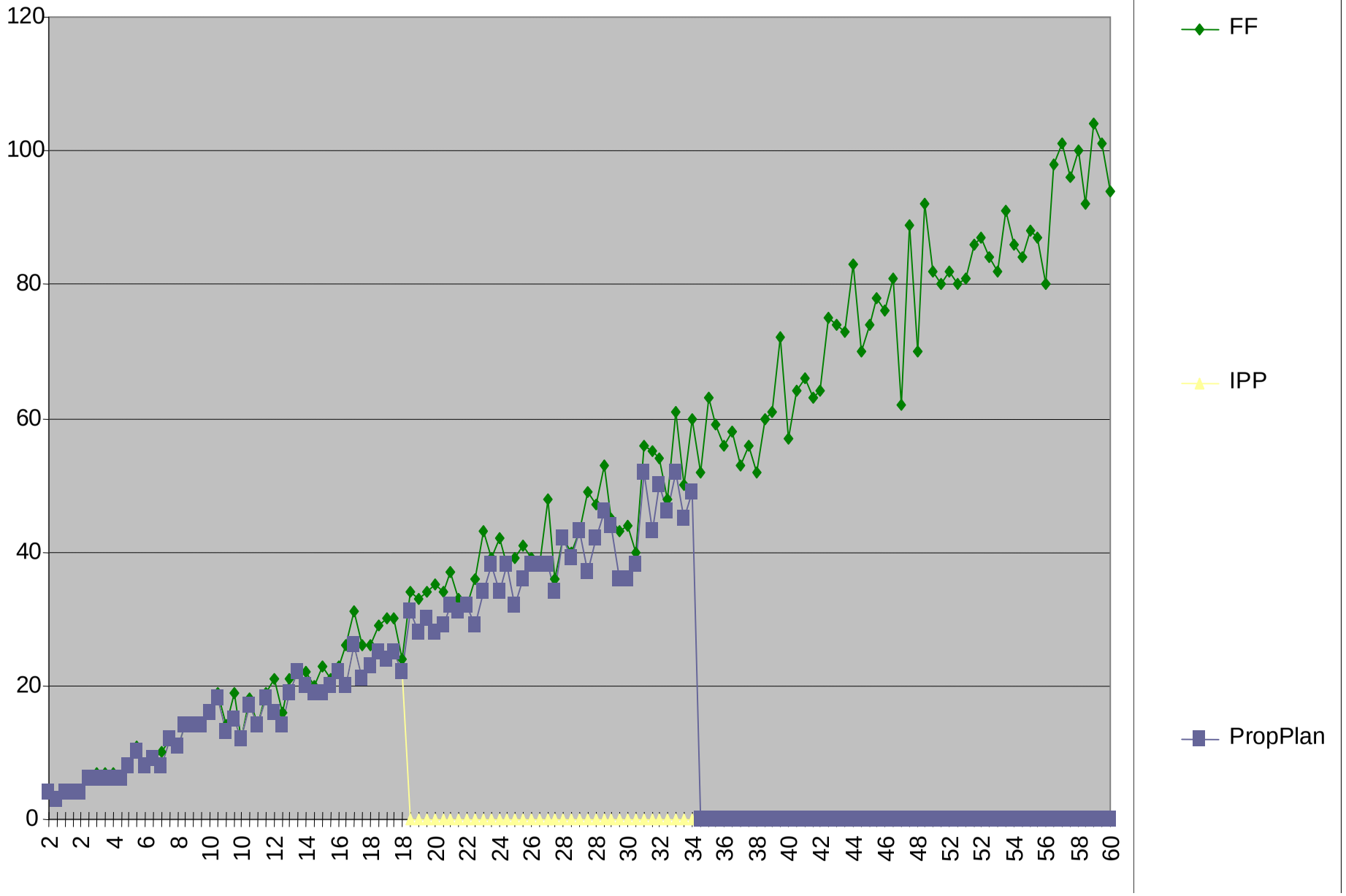
Simple Miconic 10

- HSP was the only planner to successfully solve the problems.
- Generates reasonable solutions quickly in this domain.

Full Miconic 10 Time Comparison



Full Miconic 10 # Steps Comparison

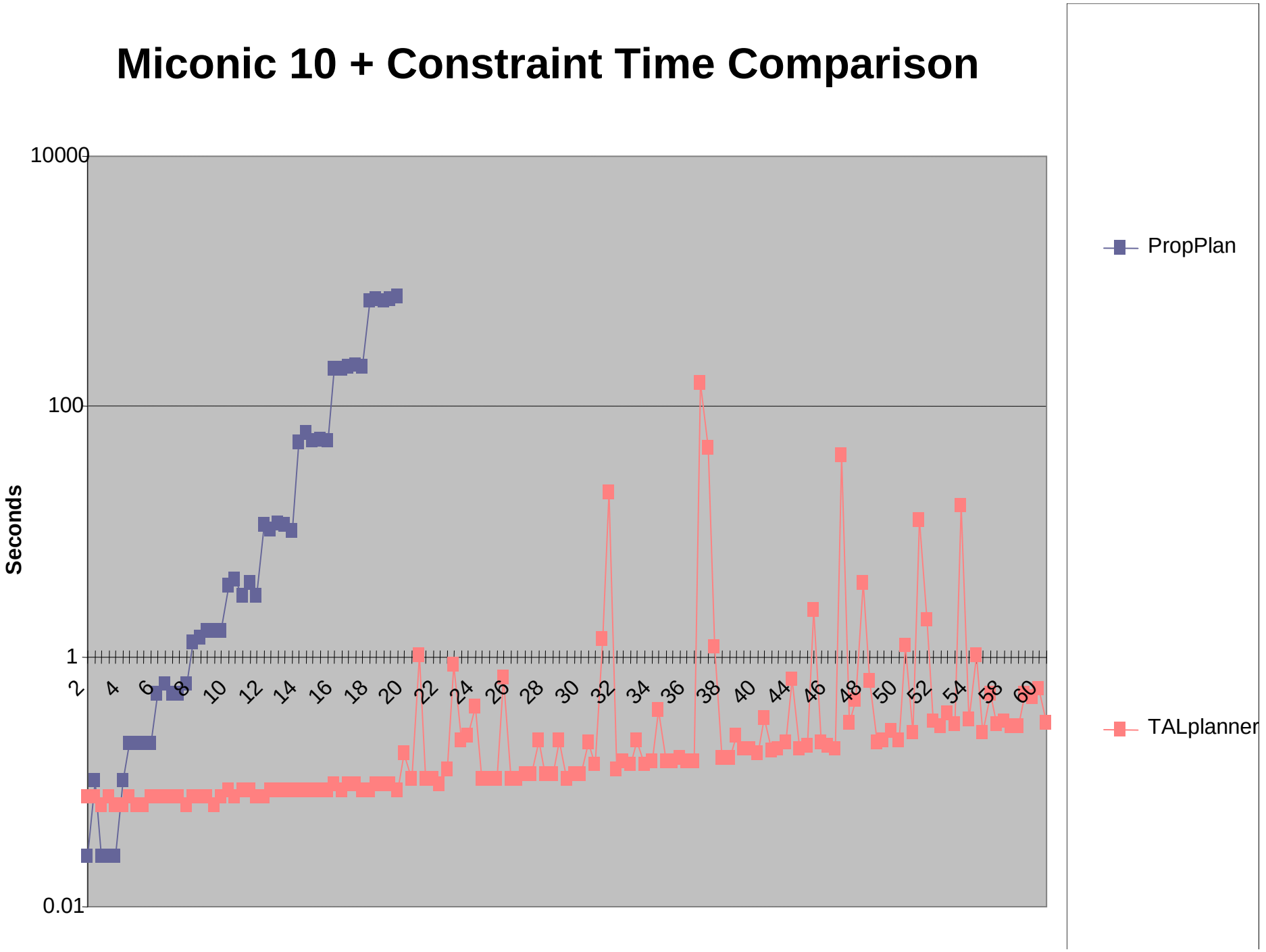


Full ADL-Miconic10

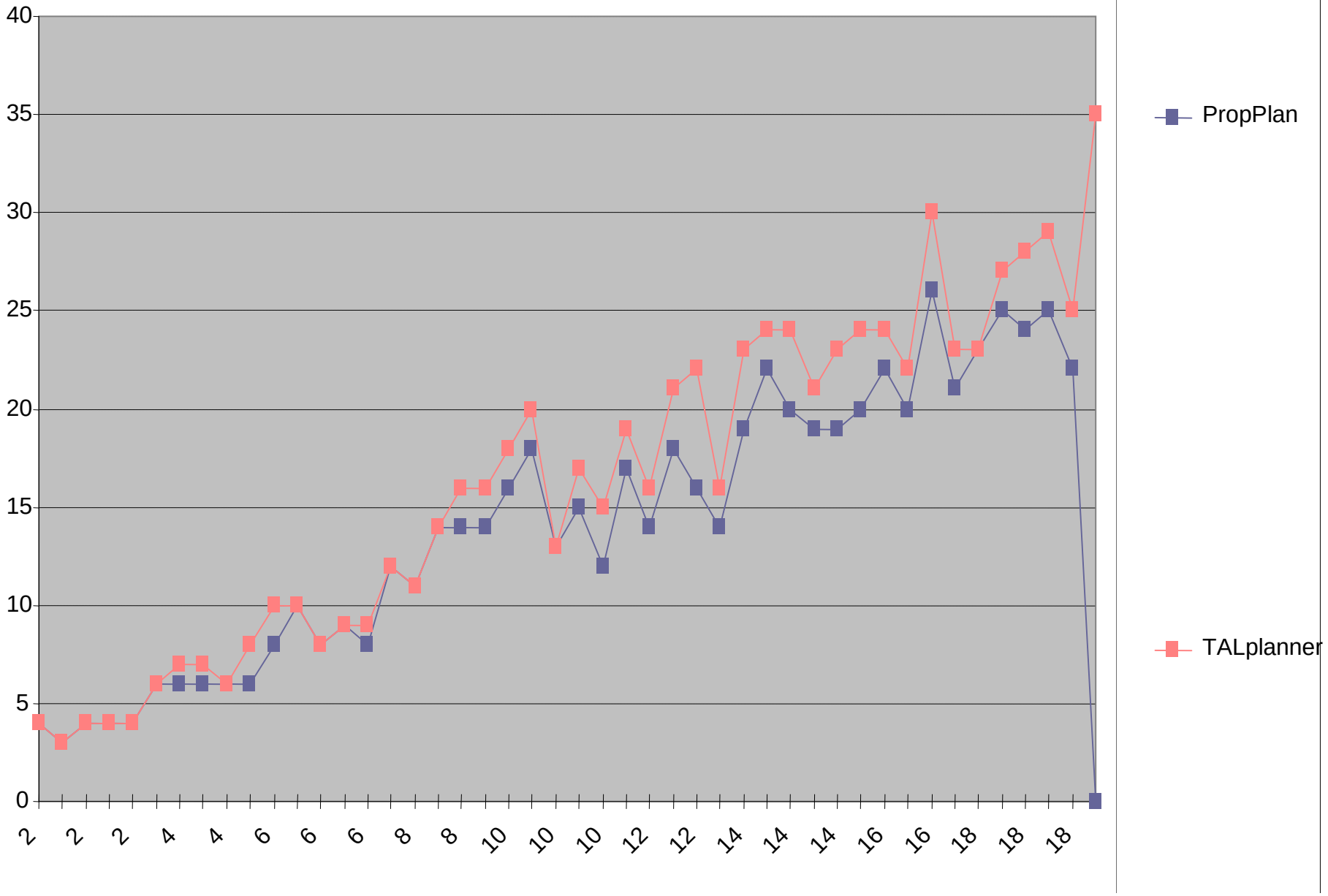
- PropPlan generates minimal length solutions.
- FF is faster.

Miconic10—ADL+Constraint

Miconic 10 + Constraint Time Comparison



Miconic 10 + Constraint # Steps Comparison



Miconic10 + Constraints

- PropPlan is generating minimal length plans.
- TALPlan is generating reasonably good plans, quickly.
- PropPlan is not using any domain specific control.

Distinguished Planners

- Schindler Lifts Ltd. Is providing a special award for performance on the Miconic-10 domain.

Distinguished Planners

- Celcorp is providing a set of awards for performance in the competition.
- There are many metrics, and it is impossible to say that any one planner was the best.
- But some planners did distinguish themselves as performing with distinction in different ways.
- I have selected a set of such planners as worthy of “special distinction.”

Distinguished Planners
Group B

- STAN
- HSP2
- MIPS
- System R

Distinguished Planners
Group A

- Two Planners demonstrated performance that was even more distinguished.
- TalPlanner
- FF