The 7th Answer Set Programming Competition

Martin Gebser, Marco Maratea, Francesco Ricca

14th International Conference on Logic Programming
and Non-monotonic Reasoning

Outline

The Seventh ASP Competition

Two years after the 6th event

- Hosted by LPNMR
- Back to the usual timeline, after the two consecutive events due to the FLoC 2014 Olympic Games

The Seventh ASP Competition

Two years after the 6th event

- Hosted by LPNMR
- Back to the usual timeline, after the two consecutive events due to the FLoC 2014 Olympic Games

Goals

- Measure the progress of the state of the art in ASP solving
- Improve benchmark suite for robust evaluation
- Study the behavior of different solving techniques
- Re-introduce a controlled form of M&S competition (tentative)

The Seventh Competition Setting

Competition Setting

- System competition and modeling competition on site
- Benchmark classification based on language features
- Benchmarks from past editions
  - $\rightarrow$ The best encodings from 2015
  - $\rightarrow$ Updated instance sets for few domains
  - $\rightarrow$ New benchmarks
- Instance selection process
- New solvers and updated versions

Outline

System Competition Format

Sub-tracks based on language features (maintained)

Track 1 (Basic) normal LP + simple built-ins

Track 2 (Advanced) + choices, aggregates, HCF disjunction, query

Track 3 (Optimization) + weak constraints

Track 4 (Unrestricted) + non-HCF disjunction

Two Categories

- Single-Processor (restricted to 1-CPU Core)
- Multi-Processor (up to 8-CPU Cores)

System Competition Format

Sub-tracks based on language features (maintained)

Track 1 (Basic) normal LP + simple built-ins

Track 2 (Advanced) + choices, aggregates, HCF disjunction, query

Track 3 (Optimization) + weak constraints

Track 4 (Unrestricted) + non-HCF disjunction

Two Categories

- Single-Processor (restricted to 1-CPU Core)
- Multi-Processor (up to 8-CPU Cores)

Marathon

- The best solver of each team
- Time limit extended by one order of magnitude

  $\rightarrow$ Assess solvers on hard instances

Setup

System Inputs

- Fixed input in ASP-Core-2
- Scripts run with fixed parameters
- Fixed encoding + instance from STD input

System Environment

- Debian Linux 64bit with Intel Xeon E5-4610v2 CPUs
- Time limits
  - Competition: 20 minutes
  - Marathon: 3 hours
- Memory Limit: 12 GB
- Multi-processor track: up to 8 cores (16 virtual CPUs)

Scoring

ASP Competition 2017 Scoring Schema

- Consider number of solved instances for decision problems
- Rank solvers on optimization problems by
  - Capability to find the optimum
  - Solution quality
- Runtime for tiebreaking

Decision, Query, Optimum found

$$\text{Score(Solver, Problem)} = \#\text{Solved(Solver)} * 5$$

Optimization Score considering solution quality

$$\text{Score(Solver, Problem)} = \sum_{\text{Instances } I} \frac{\#\text{NotBetter(Solver}, I) * 5}{\#\text{Participants}}$$

Scoring

ASP Competition 2017 Scoring Schema

- Consider number of solved instances for decision problems
- Rank solvers on optimization problems by
  - Capability to find the optimum
  - Solution quality
- Runtime for tiebreaking

Additional Criteria

- Problems are equally weighted up to 100 points each
- Incorrect answers: disqualification on per problem basis
- Final scores by summing over all problems

Benchmark Suite

A (non-)monotonic process...

Benchmarks from 2015

- Considered all the domains from 6th edition
- Selected the encoding variant that exhibited better performance in the 6th edition
- Updated instance sets for
  - Graph Colouring
- Hardness-based classification of instances
  - Inspired of SAT Competition and employed in 2015
  - Exploiting best solvers from the 6th competition
  - Robust selection done via ASP!

Benchmark Suite: Domains from past editions (T#1/2)

Domain	App	Problem	Encoding	
Graph Colouring		Decision	2014	**Track #1**
Hanoi Tower		Decision	2014	
Knight Tour with Holes		Decision	2014	
Labyrinth		Decision	2013	
Stable Marriage		Decision	2014	
Visit-all		Decision	2014	
Bottle Filling		Decision	2013	**Track #2**
Combined Configuration	√	Decision	2015	
Consistent Query Answering	√	Query	2015	
Graceful Graphs		Decision	2013	
Incremental Scheduling	√	Decision	2014	
Nomystery		Decision	2014	
Partner Units	√	Decision	2014	
Permutation Pattern Matching		Decision	2014	
Qualitative Spatial Reasoning		Decision	2014	
Reachability		Query	2013	
Ricochet Robots		Decision	2013	
Sokoban		Decision	2014	
Solitaire		Decision	2014	
Weighted-Sequence Problem		Decision	2014	

Benchmark Suite: Domains from past editions (T#3/4)

Domain	App	Problem	Encoding	
*Connected Still Life**		Optimization	2013	**Track #3**
Crossing Minimization	√	Optimization	2014	
Maximal Clique		Optimization	2014	
MaxSAT	√	Optimization	2015	
Steiner Tree	√	Optimization	2015	
System Synthesis	√	Optimization	2015	
Valves Location	√	Optimization	2013	
Video Streaming	√	Optimization	2015	
Abstract Dialectical Frameworks		Optimization	2013	**Track #4**
Complex Optimization	√	Decision	2014	
Minimal Diagnosis	√	Decision	2014	
Strategic Companies		Query	2013	

Benchmark Suite: New domains

Domain	App	Problem	
Crew Allocation	√	Decision	**Tr.#2**
Bayesian Network Learning		Optimization	**Tr.#3**
Markov Network Learning		Optimization	
Resource Allocation	√	Optimization	
Supertree Construction		Optimization	
Traveling Salesperson		Optimization	
Paracoherent Answer Sets		Optimization	**Tr.#4**
Random Disjunctive ASP		Decision	

Benchmark Classification (1)

Run the same solvers as previous classification

- clasp, lp2normal+clasp, wasp1.5
- same setting as competition
- 40 min TO (twice the timeout)

Some numbers

- 8 new domains
- about 2000 runs
- about 30 days of execution

Benchmark Classification (2)

(non-groundable) Instances that could not be grounded by any top-performing system within the timeout.

(very easy) Instances solved by all top-performing systems in less than 20 seconds.

(easy) Instances solved by all top-performing systems in less than 2 minutes.

(medium) Instances solved by all top-performing systems within the timeout.

(hard) Instances solved by at least one among the top-performing systems within 40 minutes.

(too hard) Instances that could not be solved (no solution produced in case of Optimization problems) by any of the top-performing systems within 40 minutes.

Instance Selection

Instance Selection (Criteria)

- 20 instances are included in each domain
- Exclude non-groundable and very easy instances
- Each other class shall contribute 20% to each domain
- Discard domains mostly made of easy instances
- Balance satisfiable and unsatisfiable instances for decision
- Focus on satisfiable instances for optimization
- Random selection from each class + 20% totally random

Instance Selection

Instance Selection (Criteria)

- 20 instances are included in each domain
- Exclude non-groundable and very easy instances
- Each other class shall contribute 20% to each domain
- Discard domains mostly made of easy instances
- Balance satisfiable and unsatisfiable instances for decision
- Focus on satisfiable instances for optimization
- Random selection from each class + 20% totally random

- Selection implemented in ASP!

Selection Statistics

Benchmark Suite

- 35 selected problems
- 1 too easy problem discarded
  - Resource Allocation
    → A very smart encoding by Giray Havur & Martin
- No non-groundable instances
- 700 instances selected

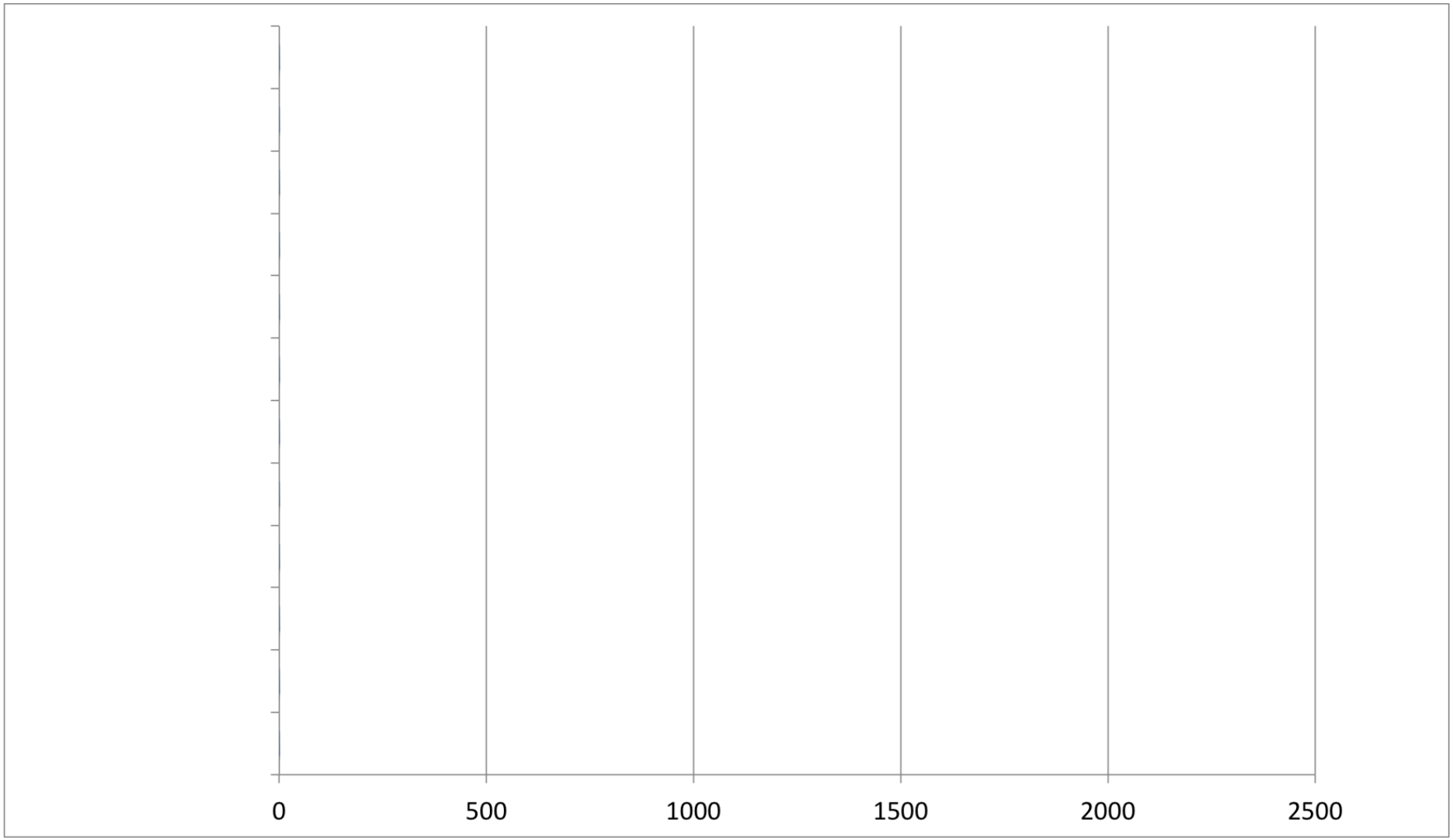
Outline

Participants

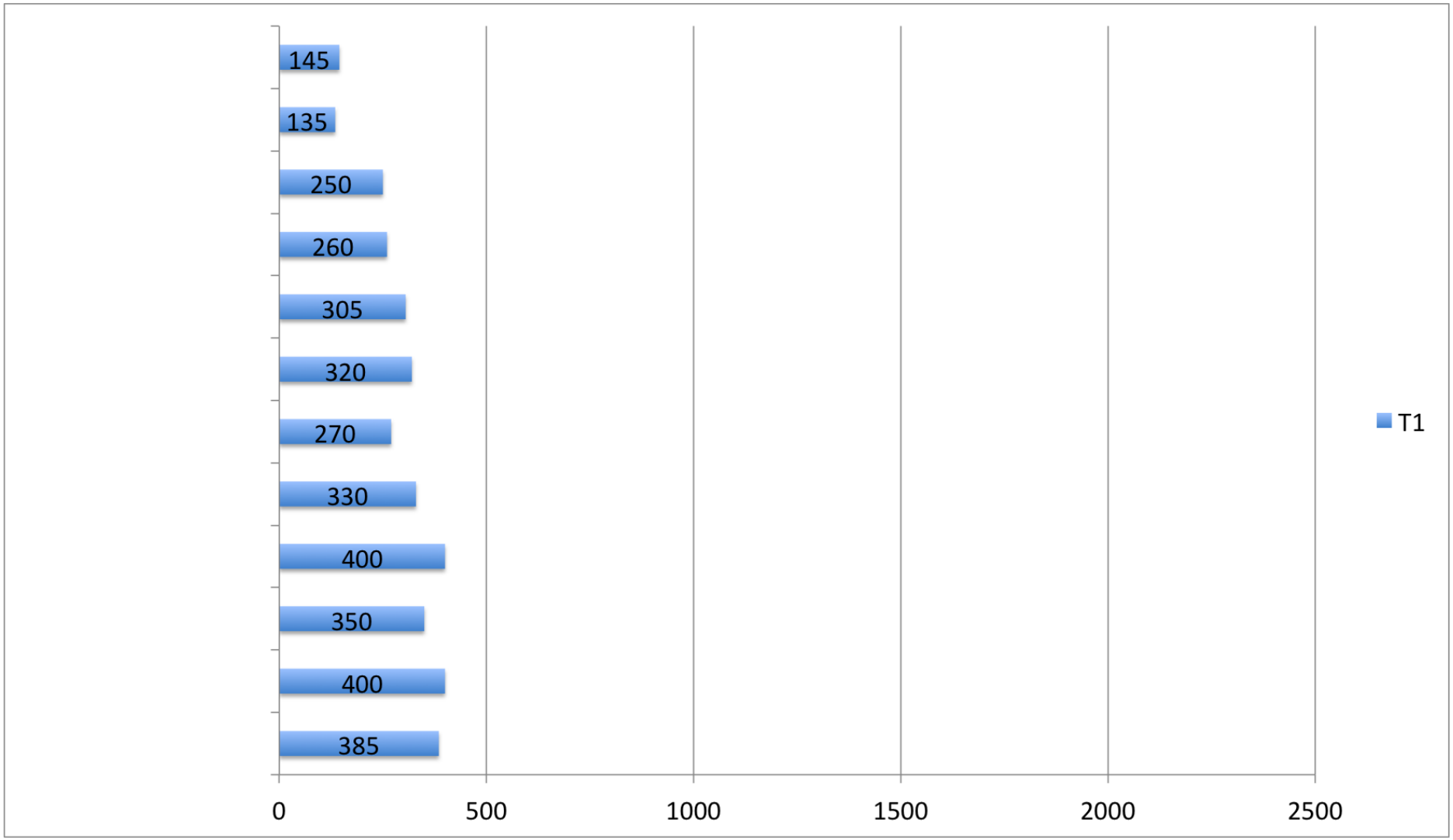
The competition featured 14 systems coming from three teams

- **Aalto Team**, Aalto University (9 solvers):
  LP2SAT+LINGELING, LP2SAT+PLINGELING-MT, LP2ACYCASP, LP2ACYCPB, LP2ACYCSAT, LP2MIP, LP2MIP-MT, LP2NORMAL, LP2NORMAL+LP2STS
- **ME-ASP Team**, University of Genoa, University of Sassari, University of Calabria (1 solver):
  ME-ASP2
- **UNICAL Team**, University of Calabria (4 solvers):
  iDLV-CLASP-DLV, iDLV+-CLASP-DLV, iDLV+-s, iDLV+-WASP-DLV

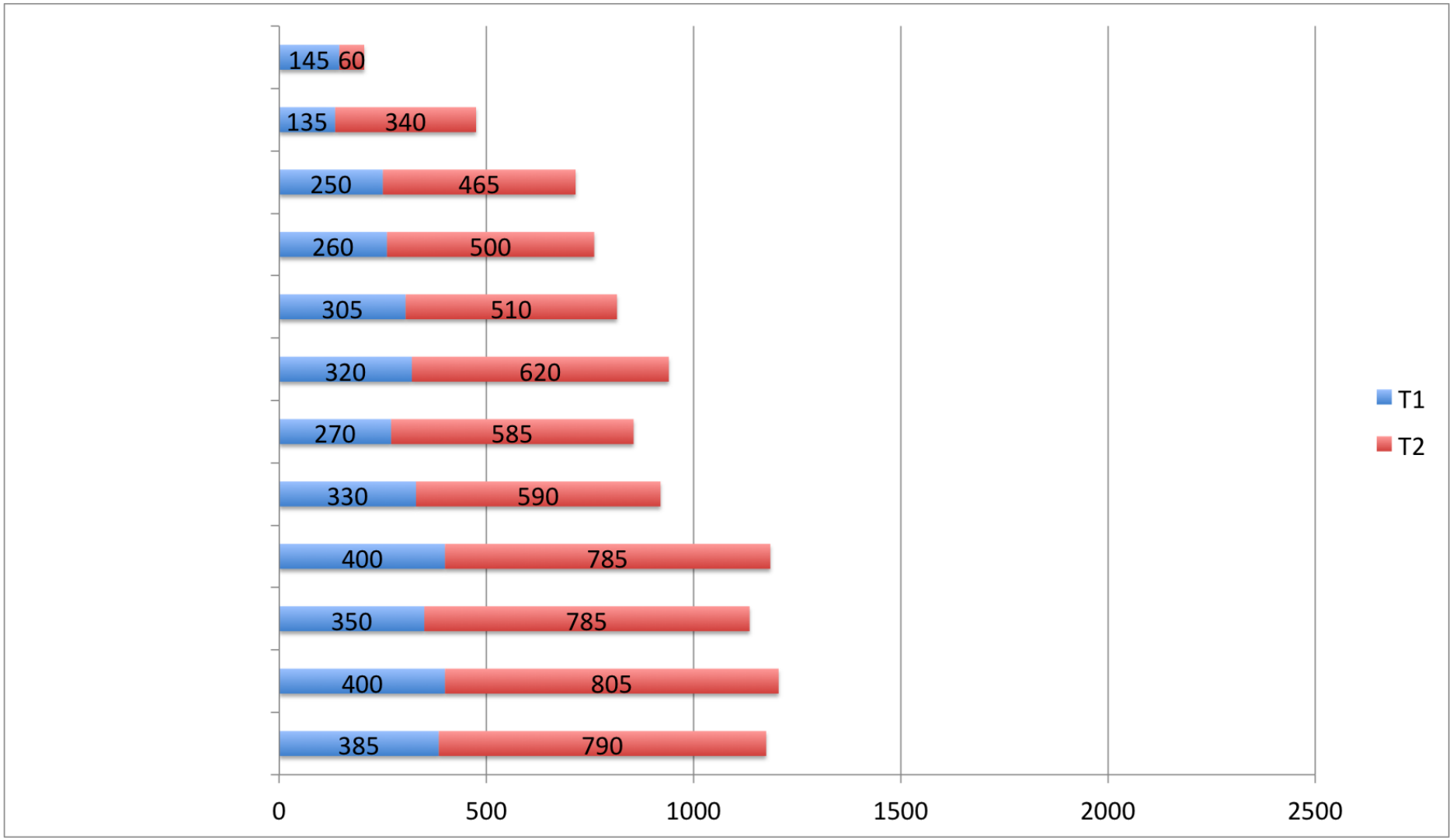
Results: Regular Solved



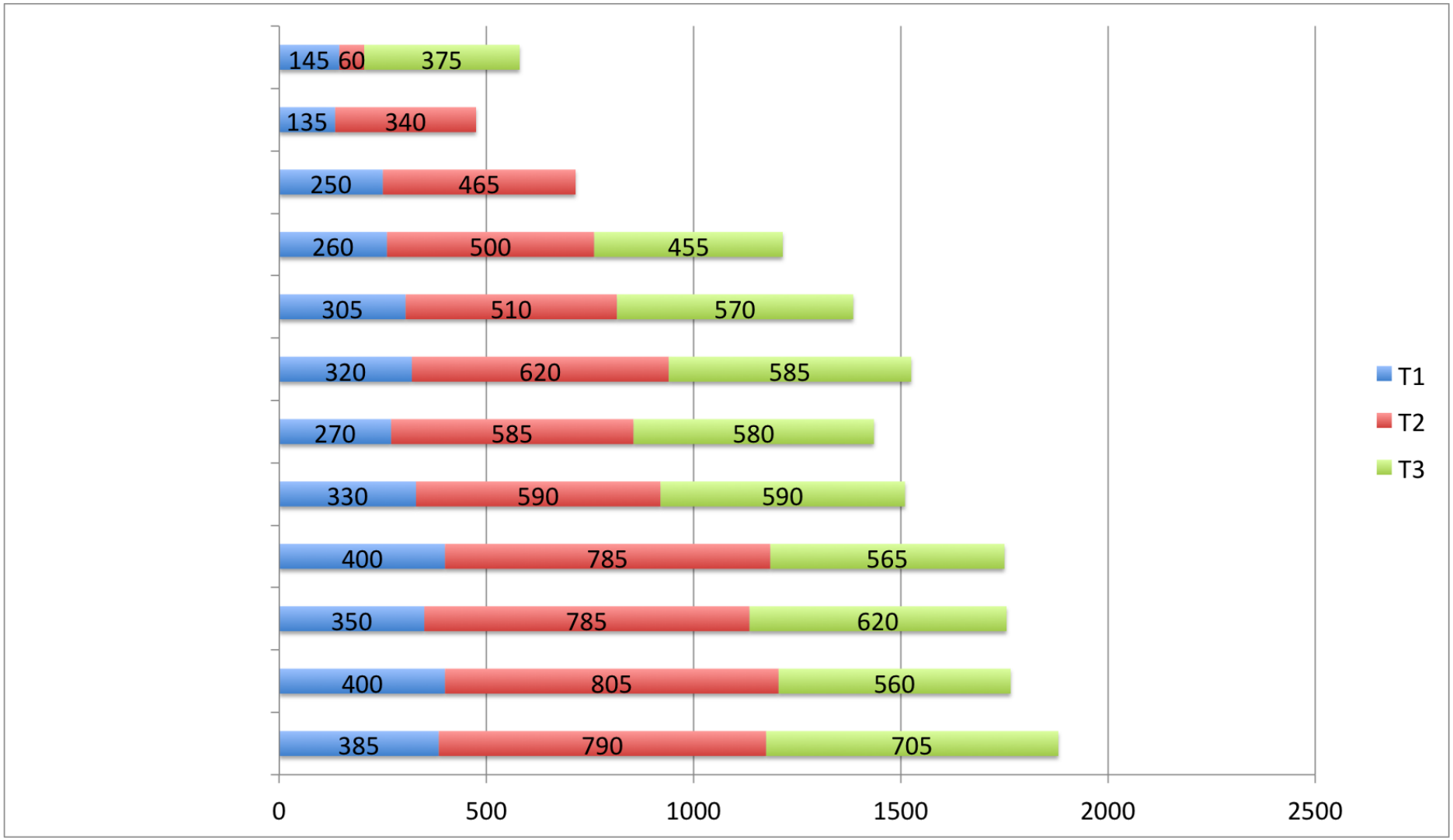
Results: Regular Solved



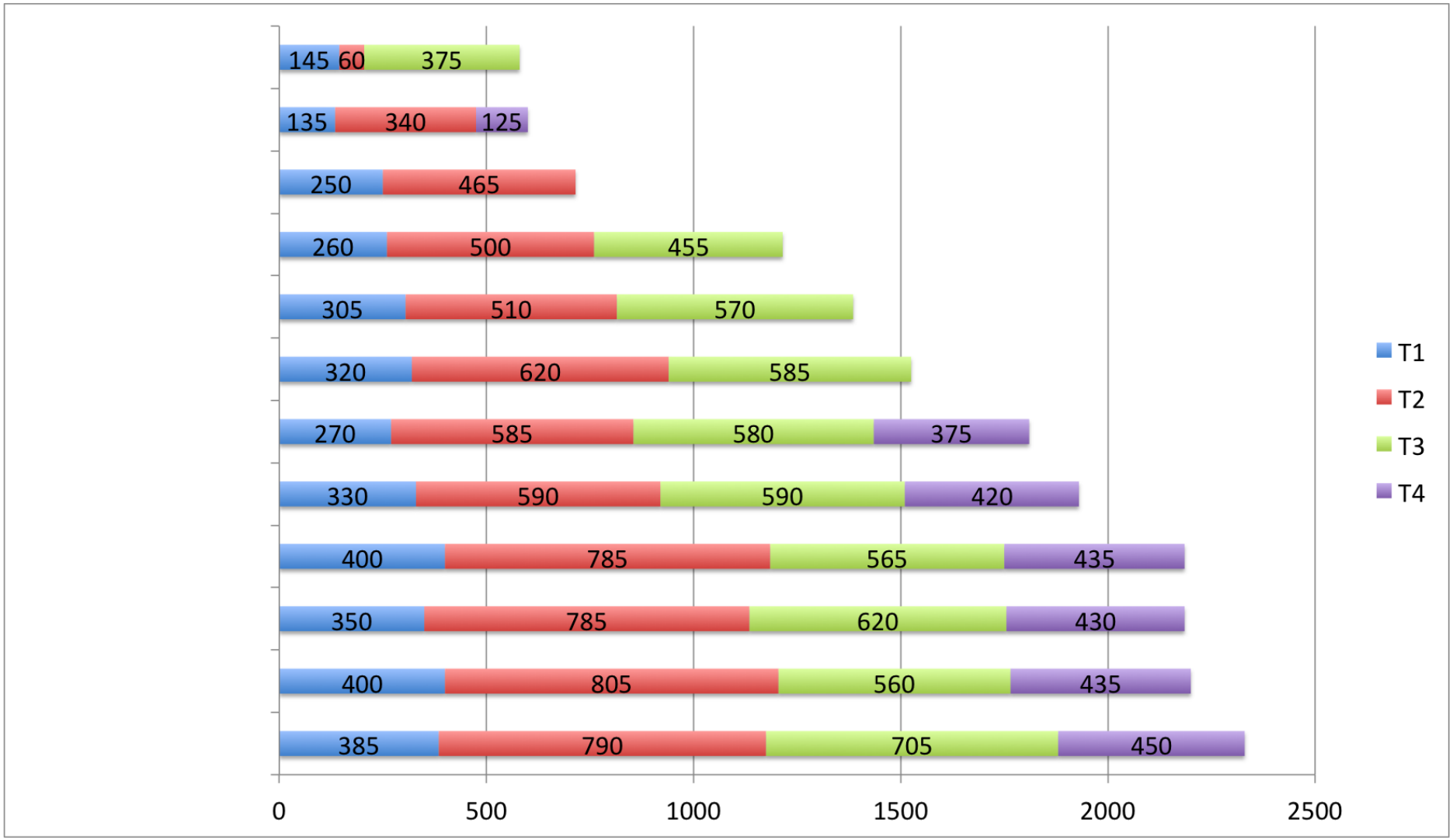
Results: Regular Solved



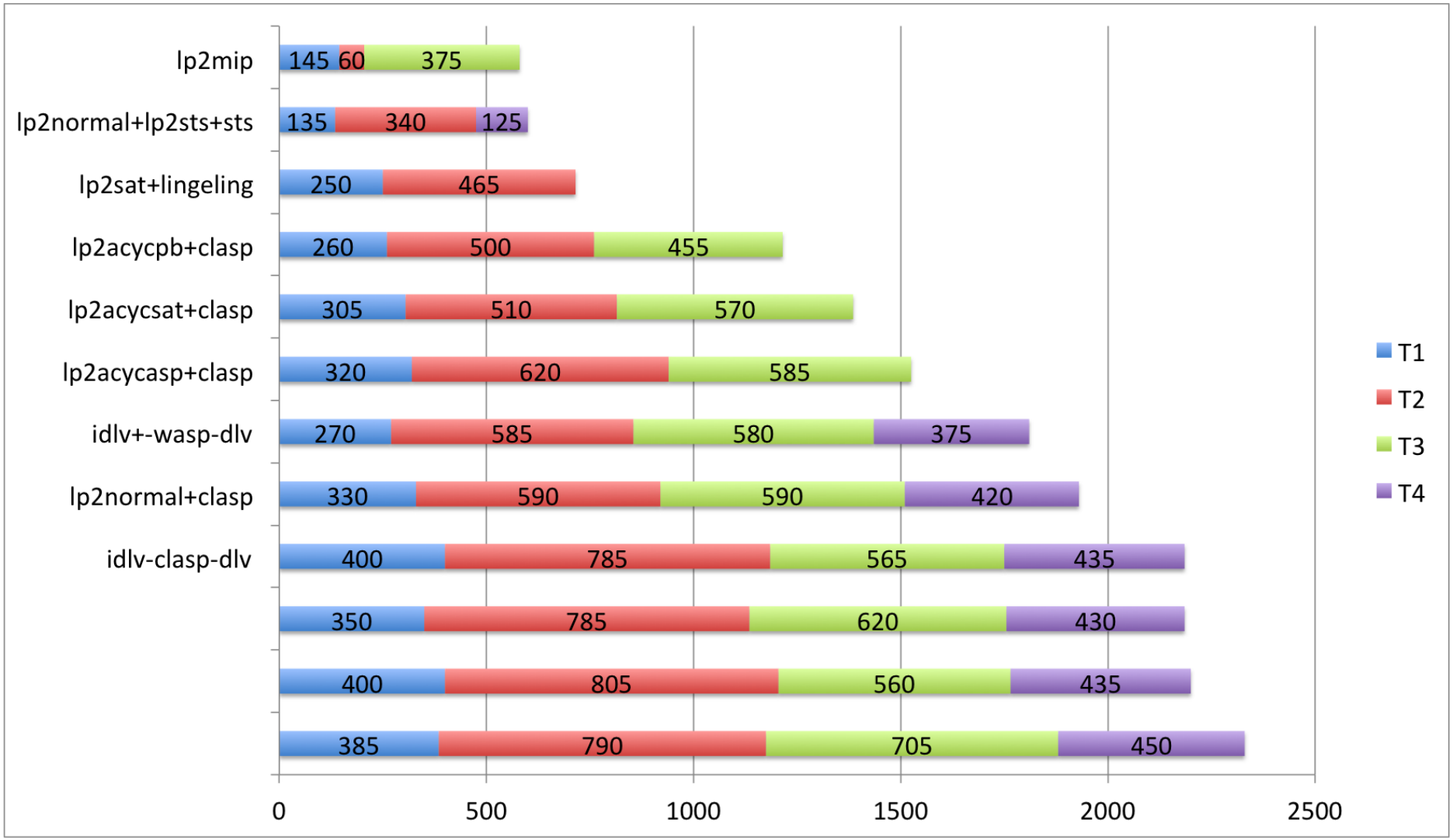
Results: Regular Solved



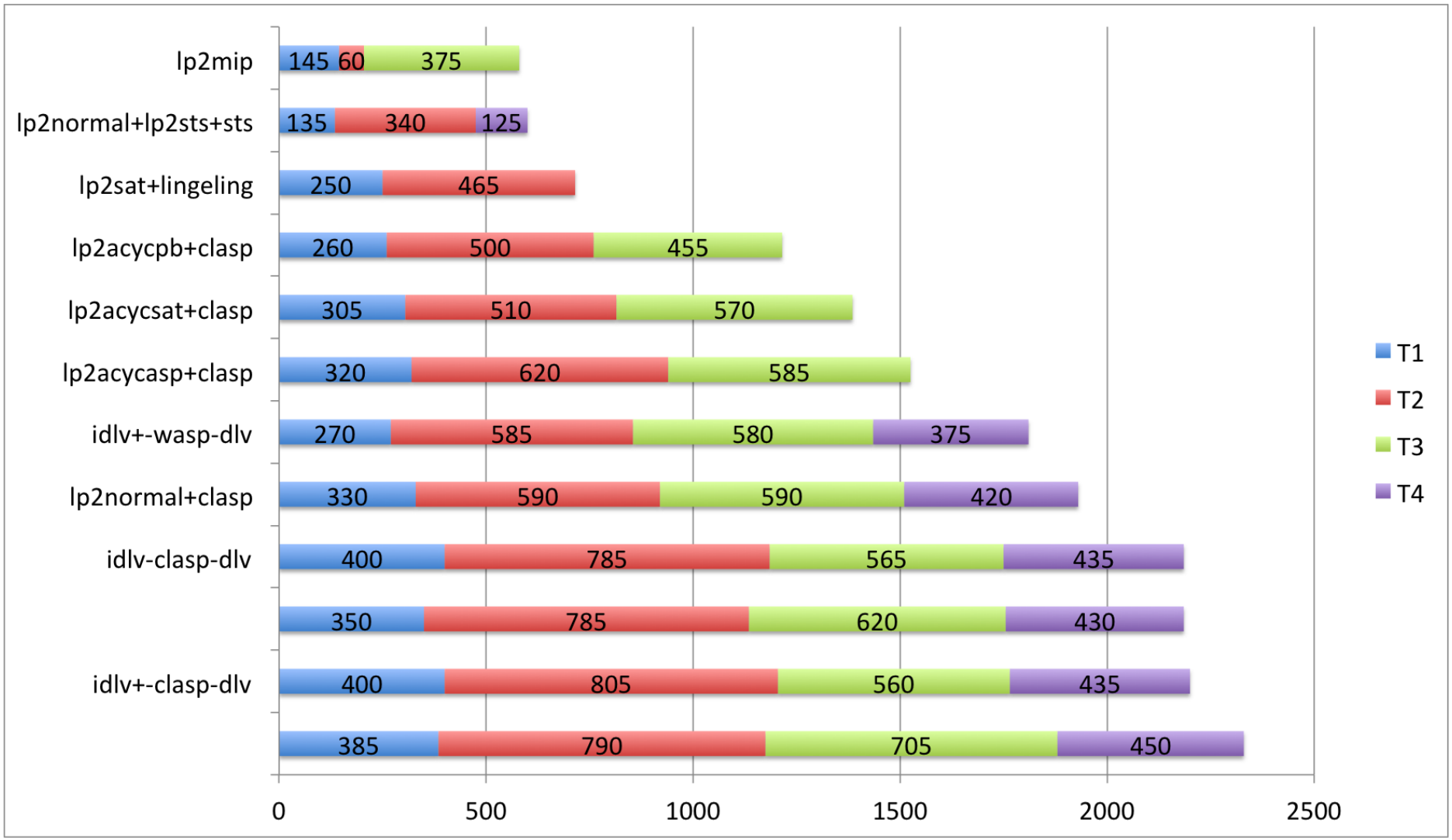
Results: Regular Solved



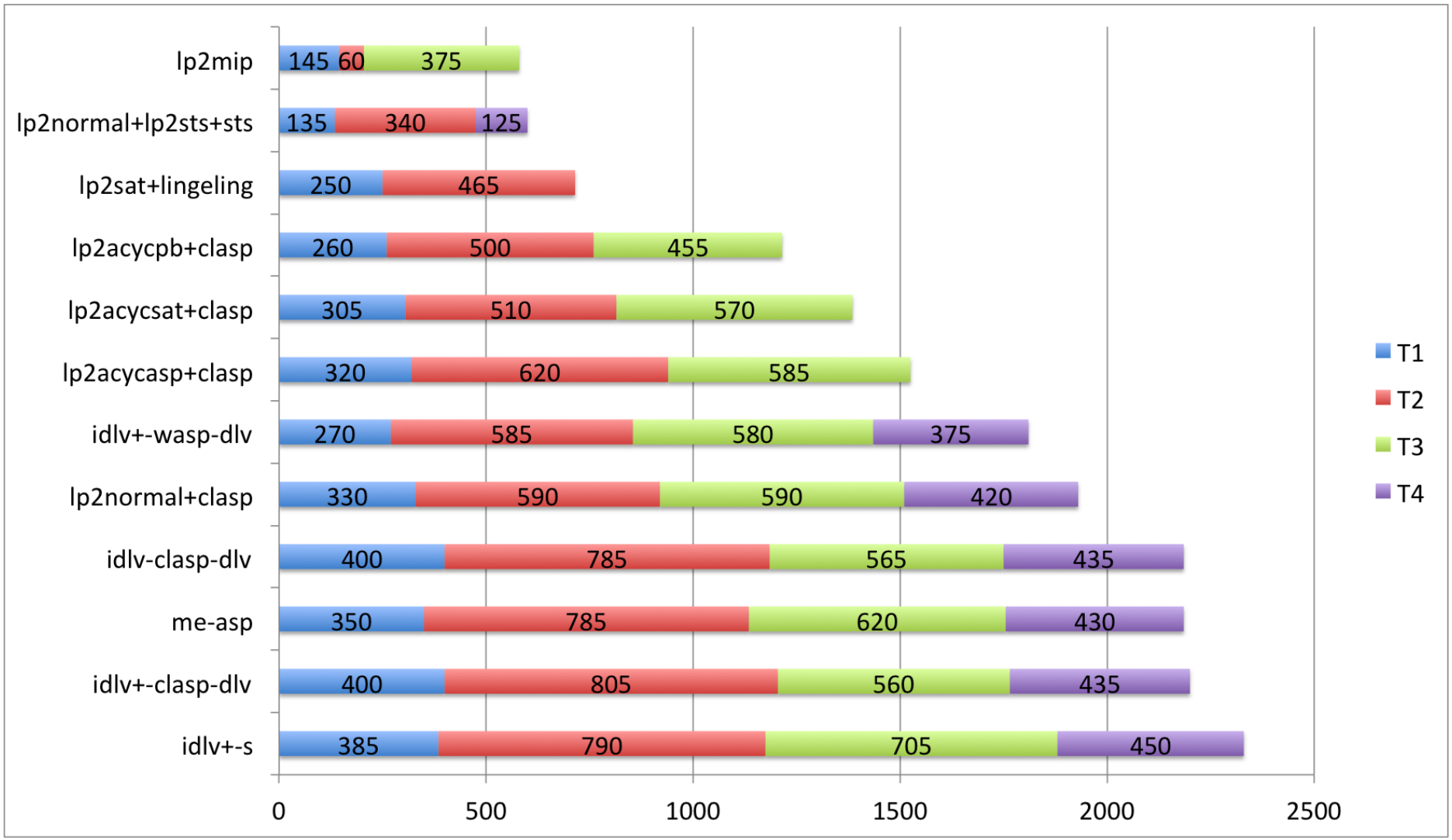
Results: Regular Solved



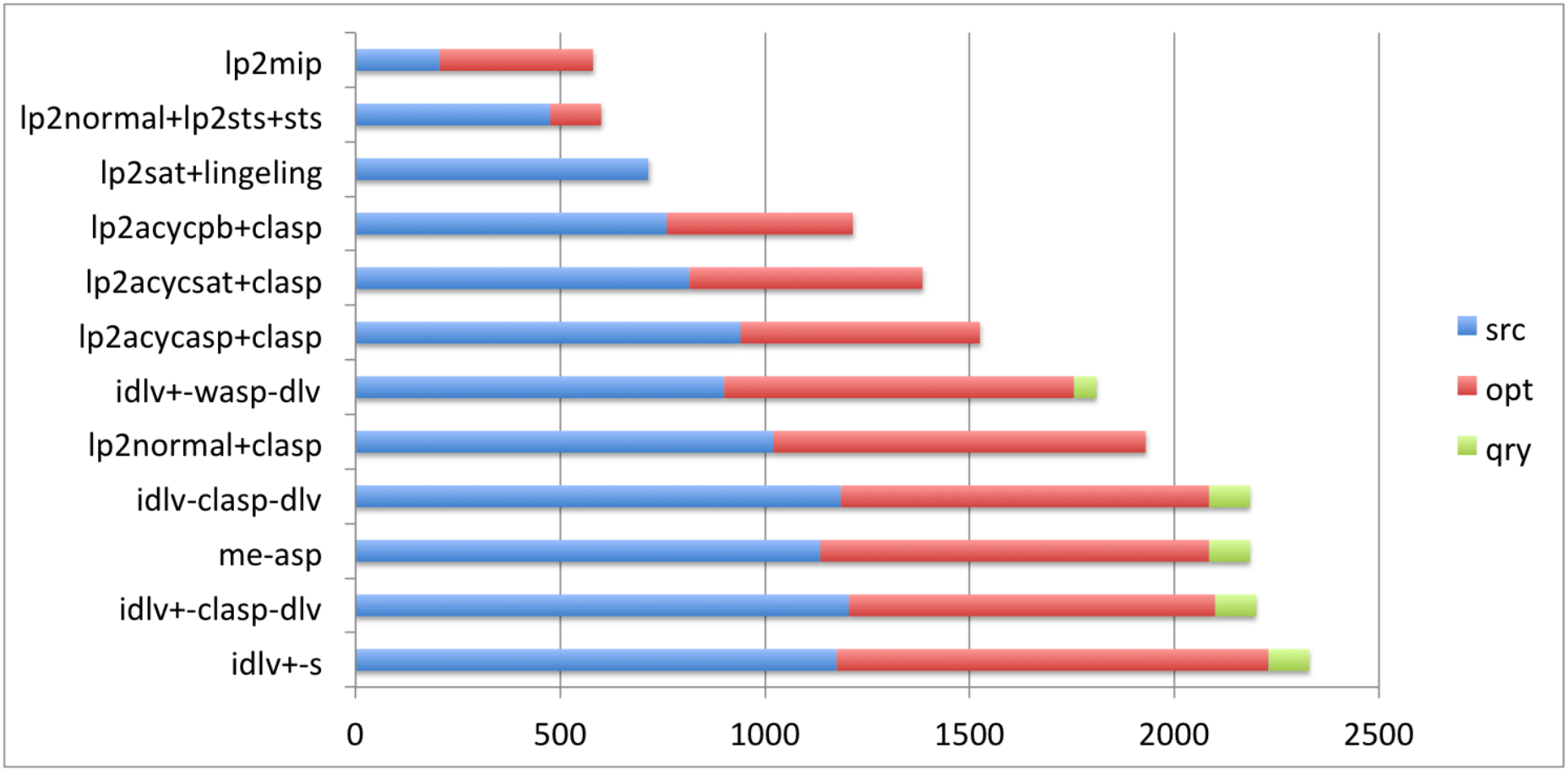
Results: Regular Solved



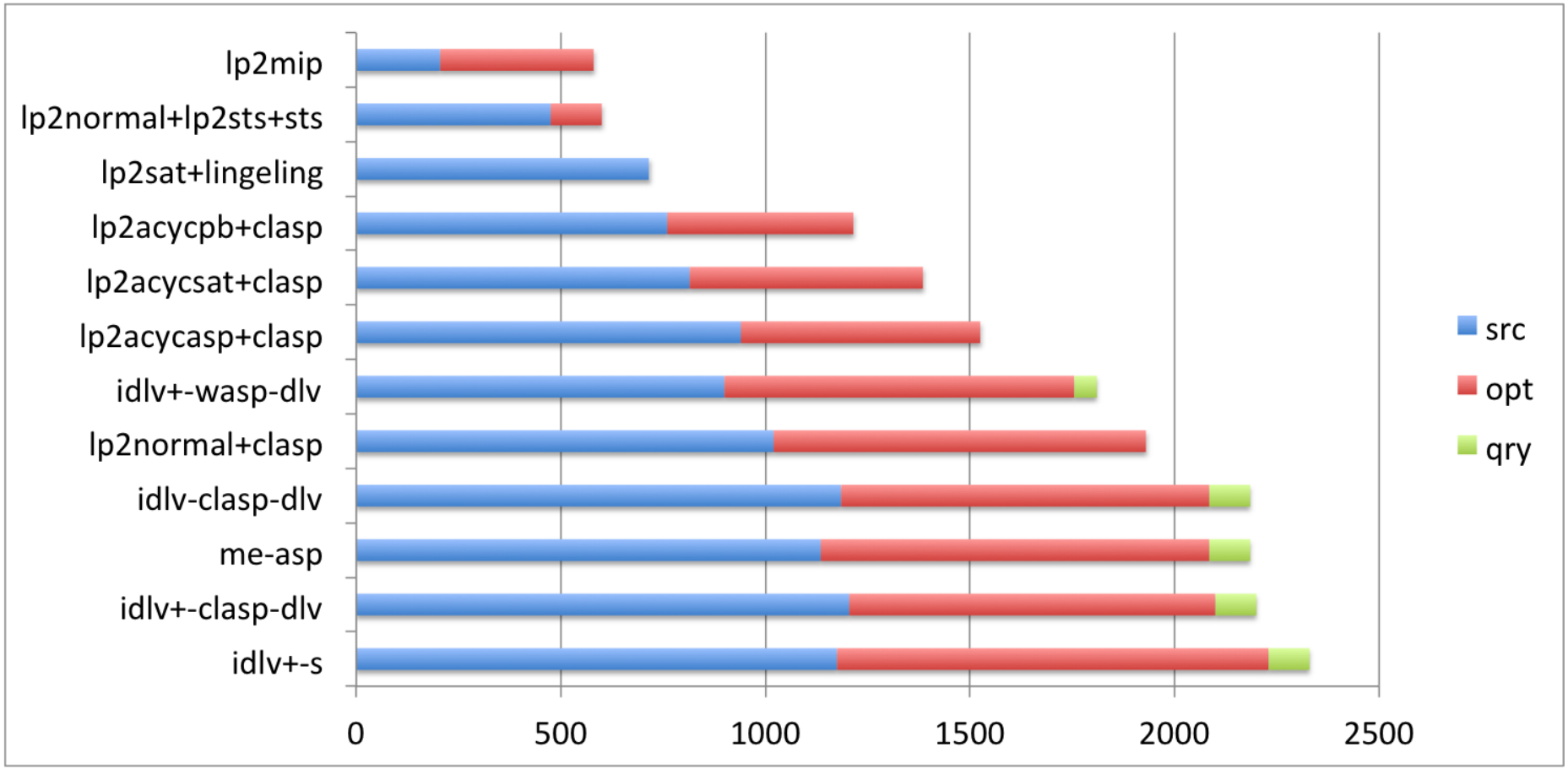
Results: Regular Solved



Results: Regular Solved By Task

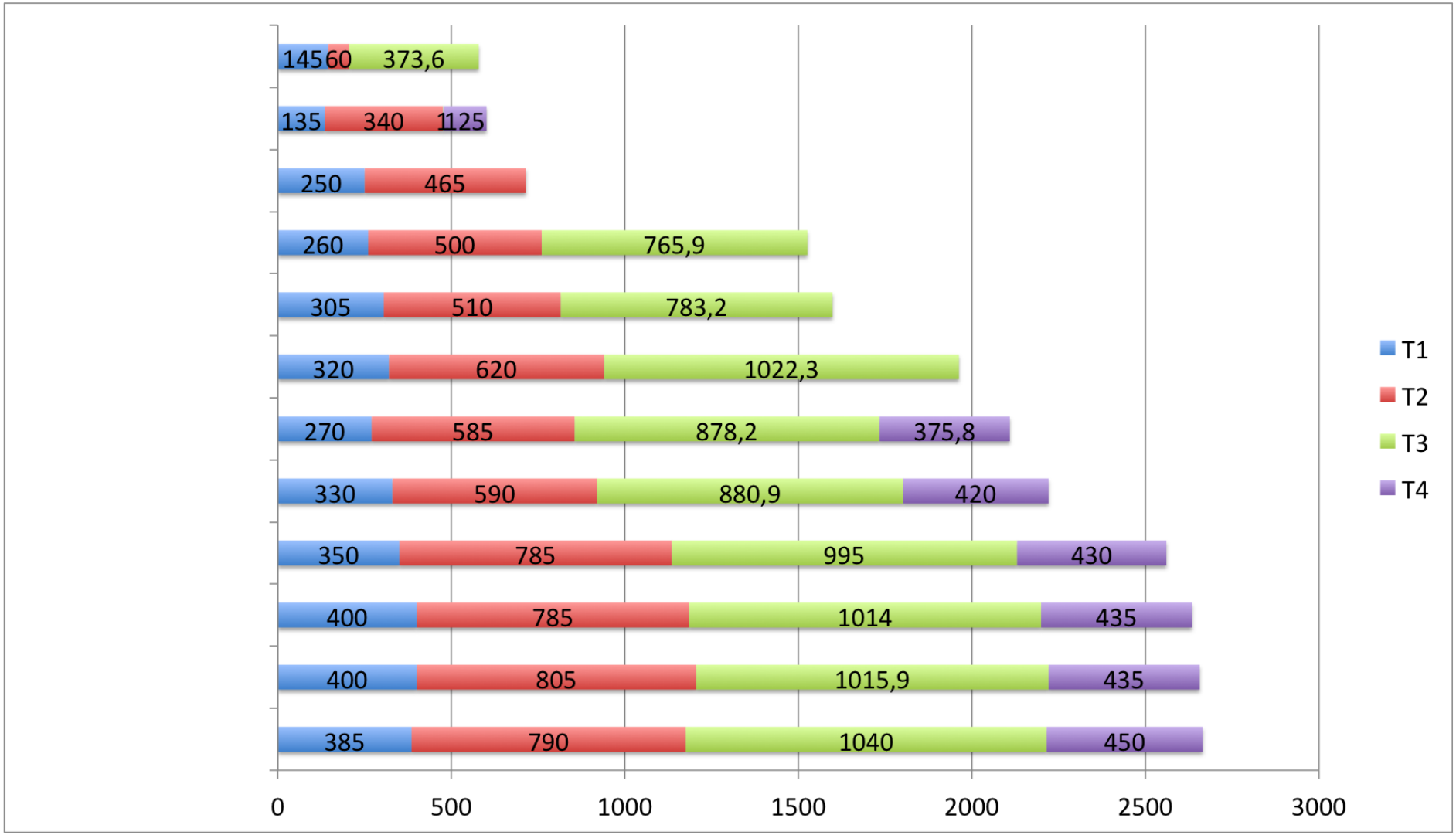


Results: Regular Solved By Task

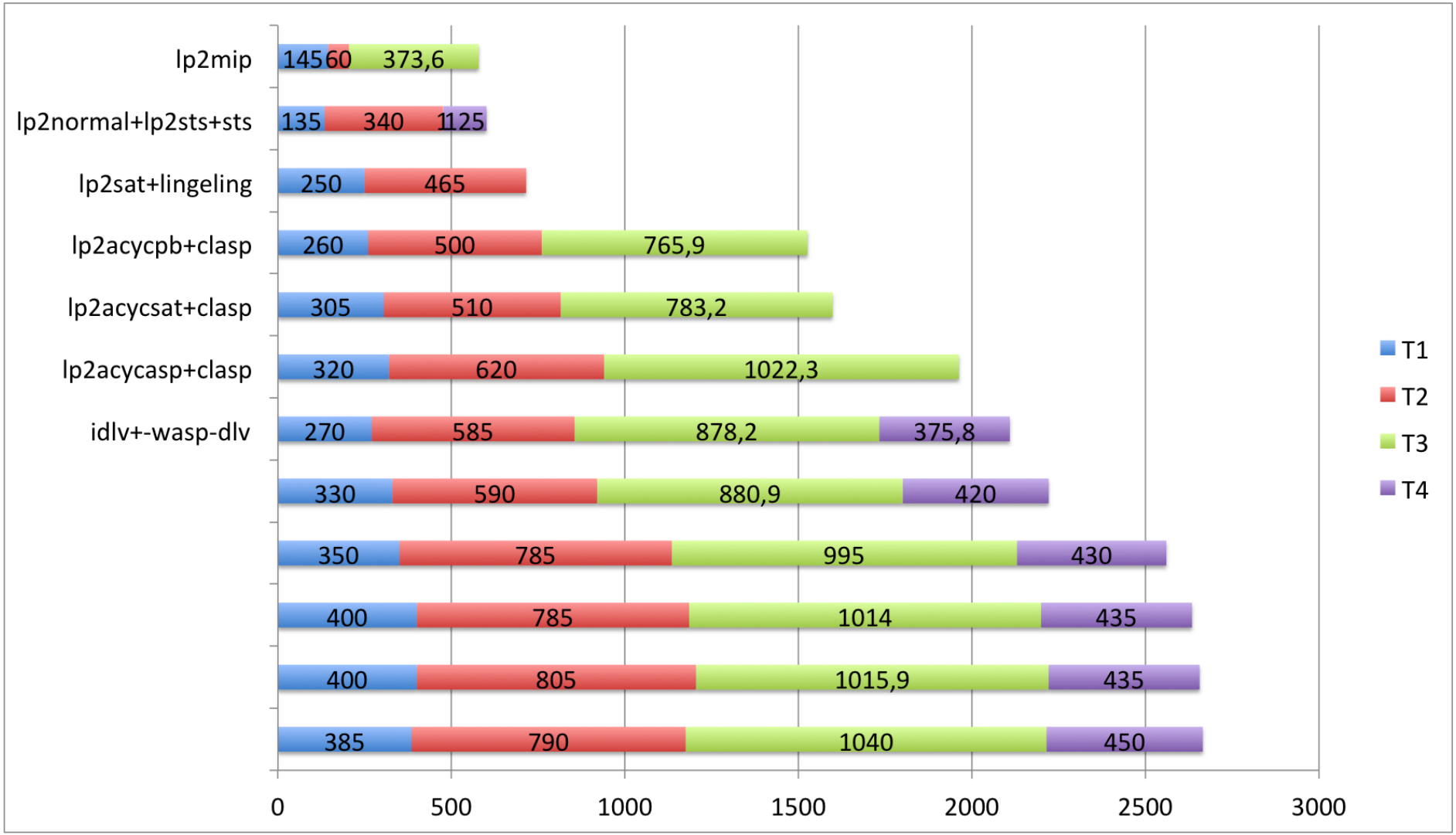


Mostly search and optimization!

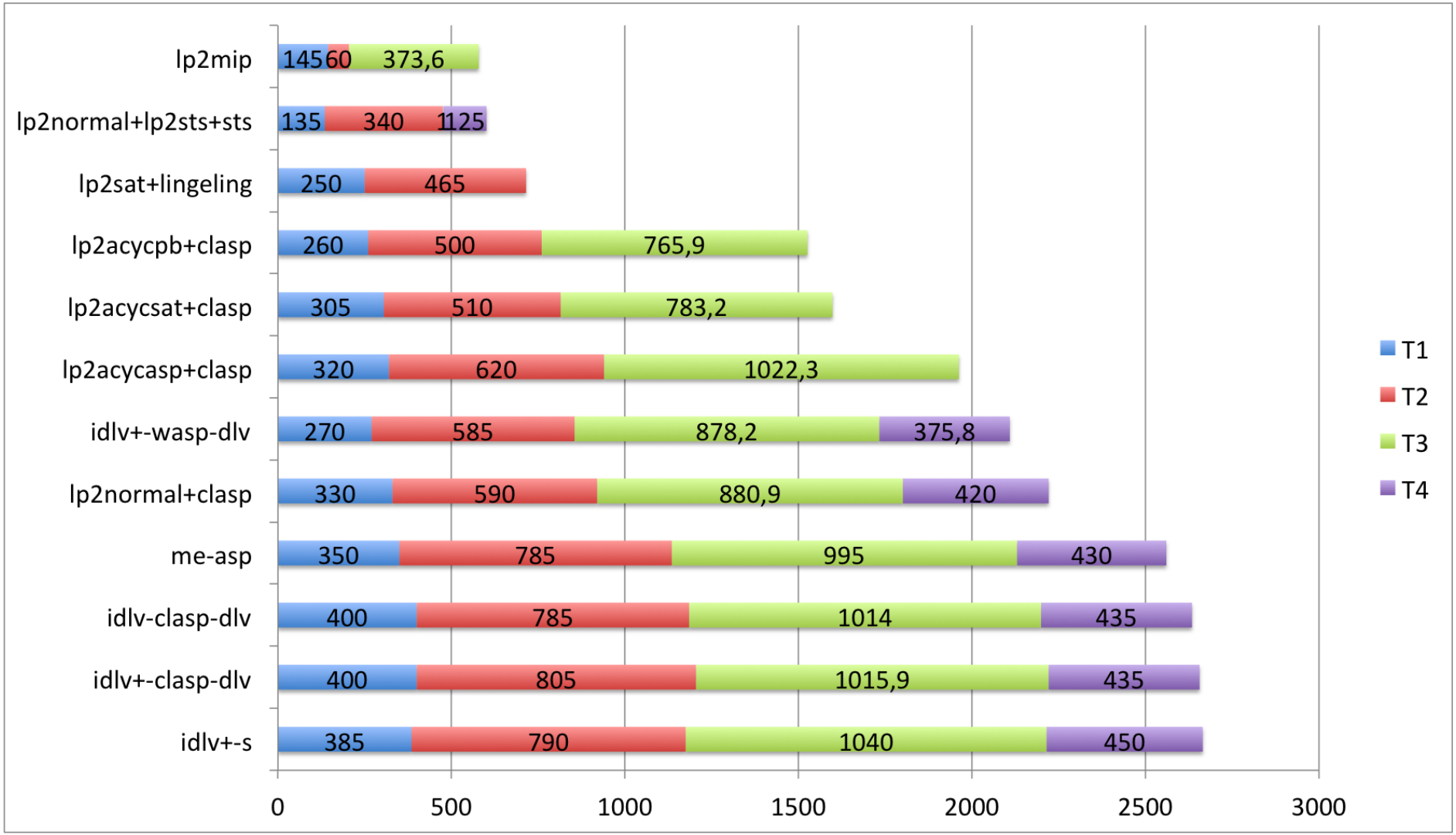
Results: By Solution Quality Score



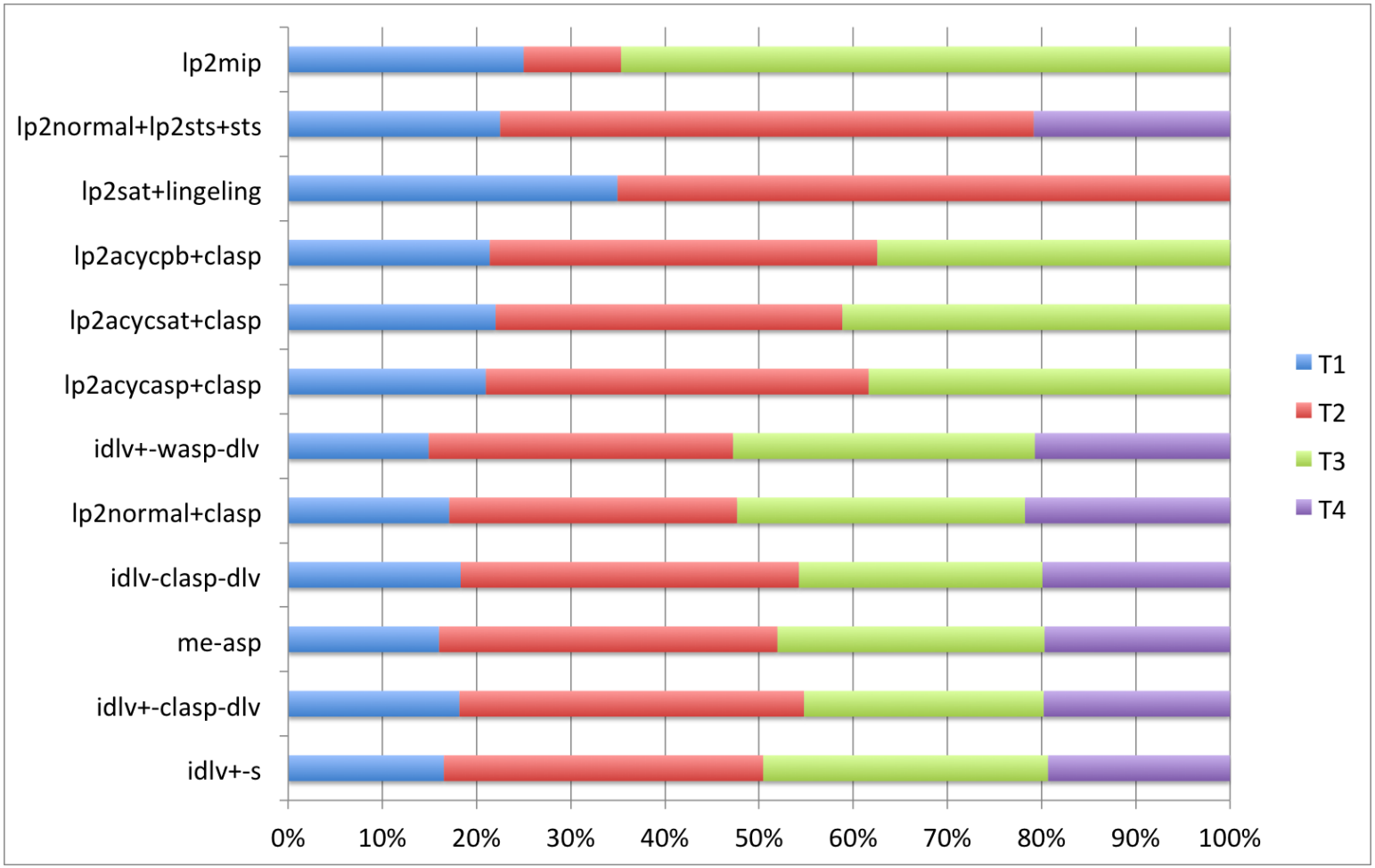
Results: By Solution Quality Score



Results: By Solution Quality Score

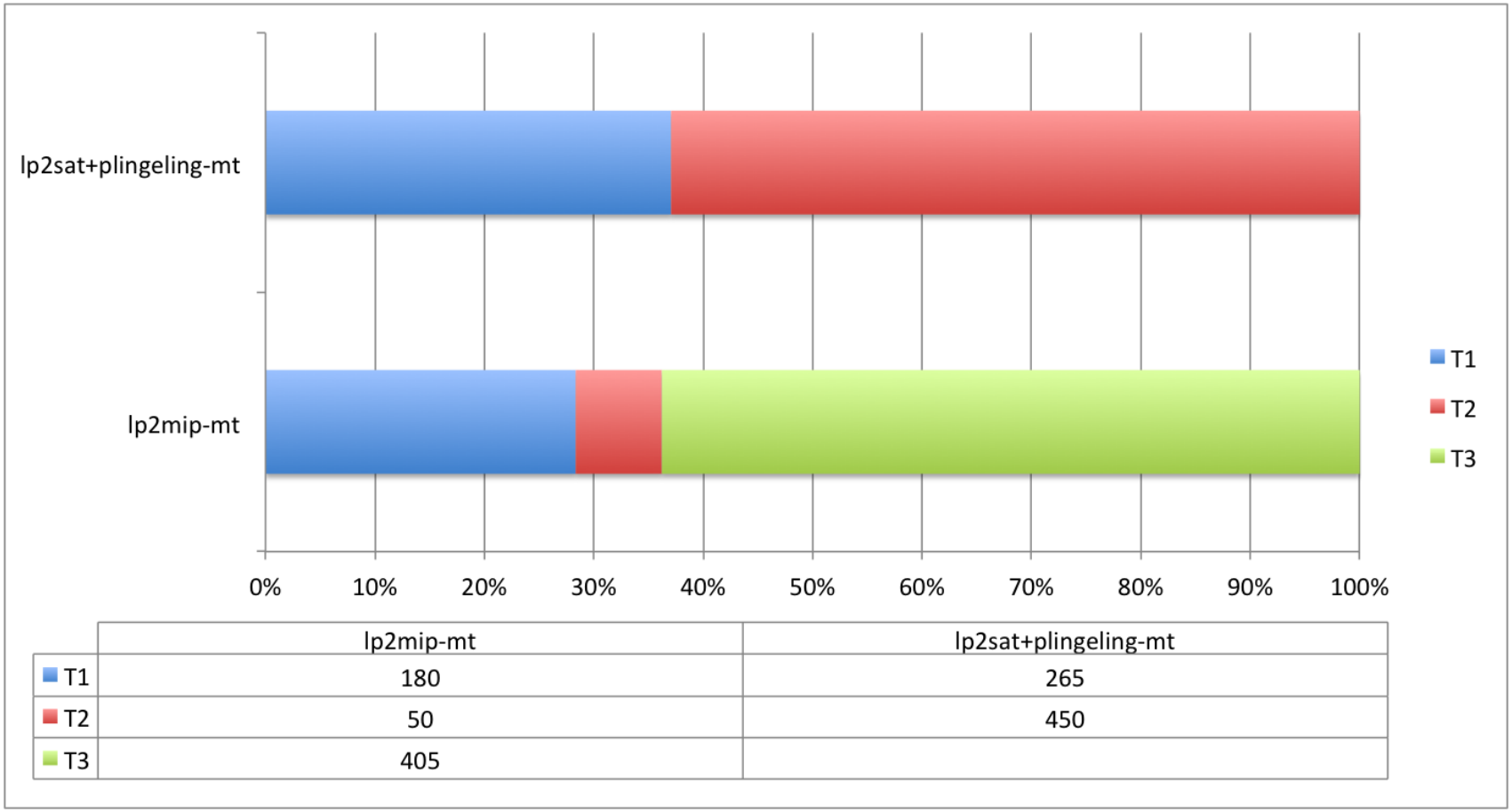


Results: Regular Percentages



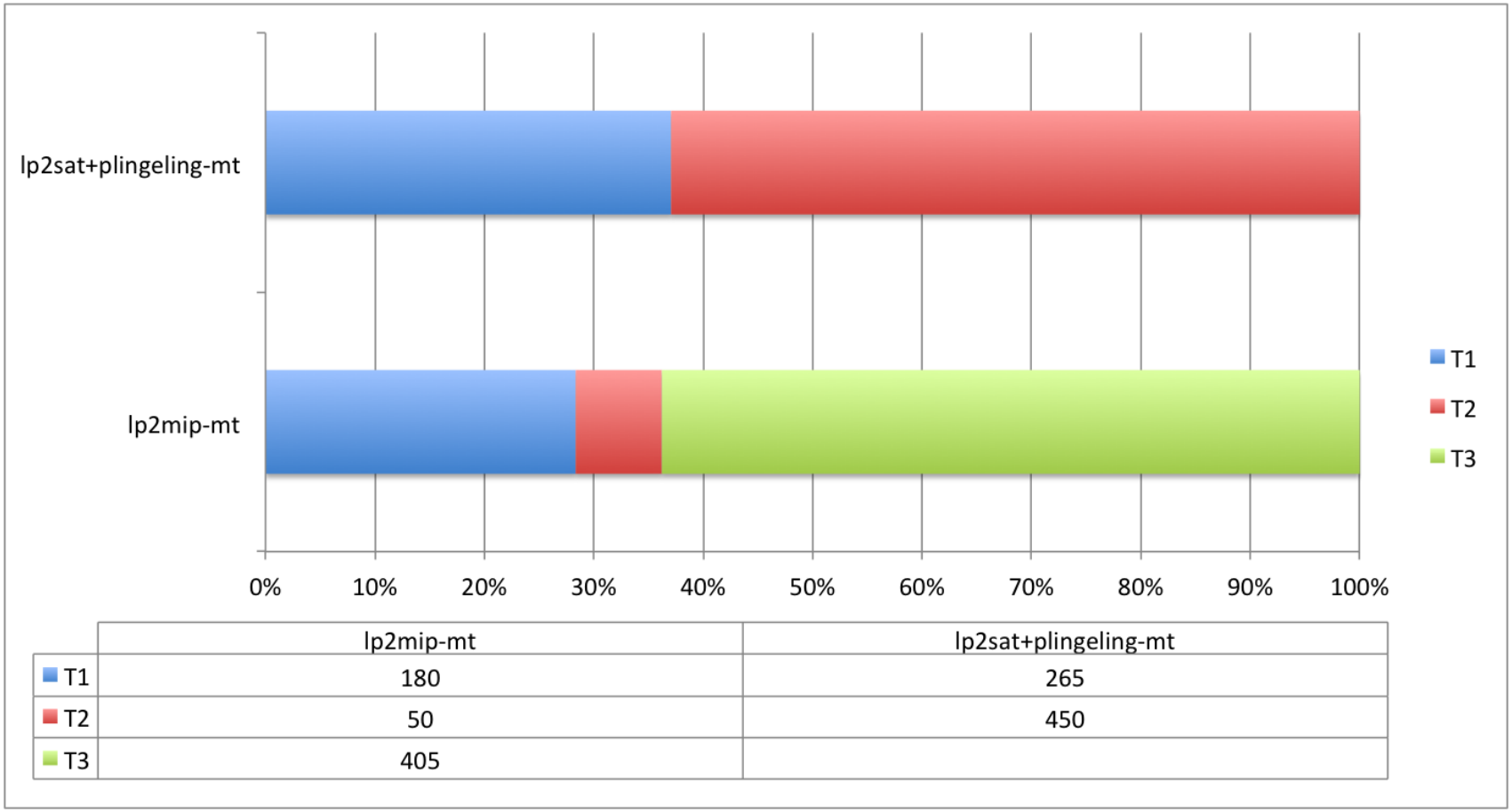
Mostly search and optimization!

Results: Multi Processor Track



	lp2mip-mt	lp2sat+plingeling-mt
T1	180	265
T2	50	450
T3	405	

Results: Multi Processor Track



	lp2mip-mt	lp2sat+plingeling-mt
T1	180	265
T2	50	450
T3	405	

lp2SAT is the winner!

Marathon Track

Marathon Track

- ...still running
- Winners will be notified next month!!

(Implemented) Suggestions from Previous Event

Scoring

- Less dependent on number of participants
- More emphasis on solved (optimal) solutions

  $\rightarrow$ 5 points is too much for non-optimal witnesses
- Two rankings?

Model and Solve Competition

- Only one domain

  $\rightarrow$ Thanks Potassco!

(Implemented) Suggestions from Previous Event

Scoring

- Less dependent on number of participants
- More emphasis on solved (optimal) solutions

  $\rightarrow$ 5 points is too much for non-optimal witnesses
- Two rankings? $\leftarrow$ DONE!

Model and Solve Competition $\leftarrow$ DONE!

- Only one domain

  $\rightarrow$ Thanks Potassco!
- Only one willing to participate!

  $\rightarrow$ Adam Smith

Suggestions for future ASP events (1)

Further Simplify Output

- Avoid using exit codes with custom semantics

  $\rightarrow$ Easy choice for SAT, not for ASP solver scripts!
- Embrace POSIX-compatible convention

  $\rightarrow$ Zero for success and non-zero for error

Benchmark Suite

- Don't stop adding ASP-oriented real-world applications
- Maintain classification by language features
- Maintain some more easy domains

Suggestions for future ASP events (2)

Reasoning Tasks

- Brave Reasoning
- Answer Set Counting
- Extended language features

Extend the ASP Development community

- ASPLib web site
- Lower the entrance barrier for newcomers

Offline Modeling Competition

- Change the competition format
- Make it a competition for students
- Involve some big company
- Have a prize in money

Suggestions for future ASP events (3)

Suggestions from the PAoASP panel

- More reasoning tasks: counting, etc.
- Grounding/Lazy grounding
- Continuous evaluation
- Emphasize knowledge representation
- Extend the standard $\rightarrow$ New Tracks
- Starexec

Suggestions for future ASP events (3)

Suggestions from the PAoASP panel

- More reasoning tasks: counting, etc.
- Grounding/Lazy grounding
- Continuous evaluation
- Emphasize knowledge representation
- Extend the standard $\rightarrow$ New Tracks
- Starexec

Further suggestions

- Fresh (wo)men power

Suggestions for future ASP events (3)

Suggestions from the PAoASP panel

- More reasoning tasks: counting, etc.
- Grounding/Lazy grounding
- Continuous evaluation
- Emphasize knowledge representation
- Extend the standard $\rightarrow$ New Tracks
- Starexec

Further suggestions

- Fresh (wo)men power
- Different structure of the OC

  $\rightarrow$ Easier to pass the relay baton

Thanks

We would like to thank all members of the community for the support

Thanks

We would like to thank all members of the community for the support

Finally, do not forget award ceremony!