



Università degli studi di Padova

Dipartimento di Tecnica e Gestione dei Sistemi Industriali

Motion Planning vs. Automated Planning in benchmarking

M. Reggiani E. Pagello

Open Forum on Good Experimental Methodology and
Benchmarking in Robotics

Once upon a time...



- ... in a land (not) so far away, full of bicycles and canals, a group of people met together to speak about motion planning benchmarks



Meeting on motion planning benchmarks

March 12th, 2004

before the EURON annual meeting in Amsterdam

partially funded by an Euron Special Study on Motion Planning Benchmarks

Four years ago scenario in MP



- Many important contributions, robust and efficient tools plan collision-free paths for many d.o.f. robots in complex environments;
- No open, standardized benchmarks and performance assessment methodology;
- Each tool measured by dedicated set of problems;
- No sharing of problems due to proprietary formats.



Current scenario in MP



- Many important contributions, robust and efficient tools plan collision-free paths for many d.o.f. robots in complex environments;
- No open, standardized benchmarks and performance assessment methodology;
- Each tool measured by dedicated set of problems;
- No sharing of problems due to proprietary formats.



Is it really the current scenario?



- 1023 pages, “benchmark” word used 6 times

Algorithms for Planning and Control of Robot Motion
Main :: Home Page

Edit History Print

- ◆ Home
- ◆ News
- ◆ Events
- ◆ Literature
- ◆ PhD Theses
- ◆ Success Stories
- ◆ Video Showcase

IEEE RAS Technical Committee on Algorithms for Planning and Control of Robot Motion

Scope
As modern robots address real-world problems in dynamic, unstructured, and open environments, novel challenges arise in the areas of robot control algorithms and motion planning. These challenges stem from an increased need for autonomy and flexibility in robot motion and task...

- benchmark?

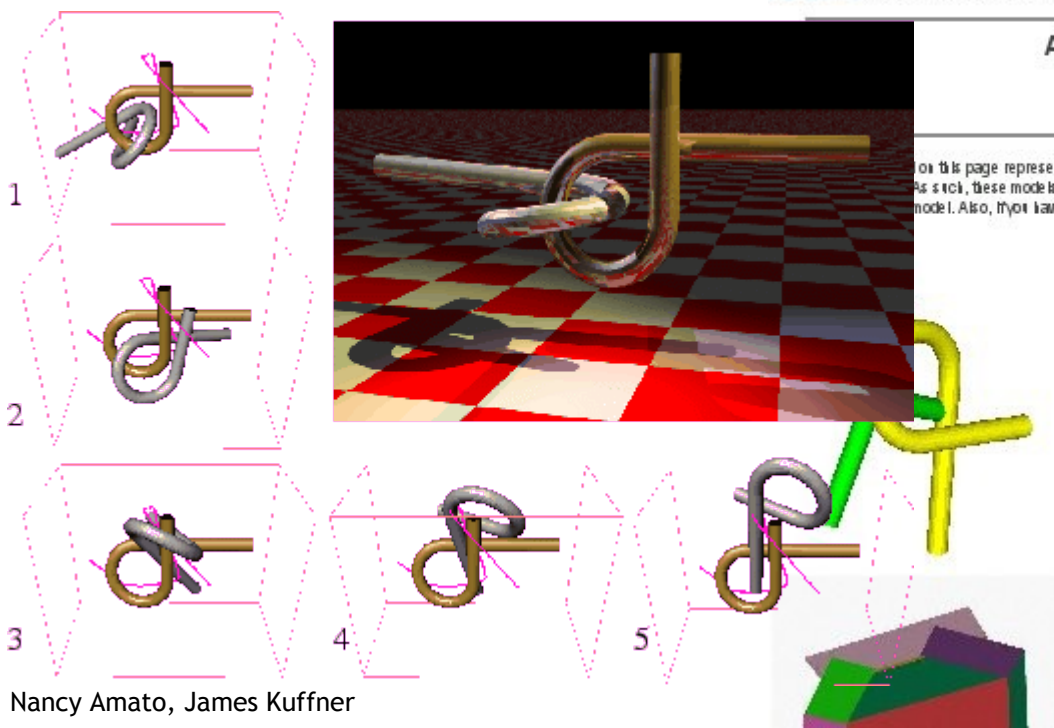


Is it really the current scenario?

- Google search....



Alg & App Group | Home | Research | Publications | People | Resources | News



Nancy Amato, James Kuffner

Algorithms & Applications Group Motion Planning Puzzles (aka benchmarks)

On this page, we represent our attempt to try to start a collection of benchmark problems that can be used to compare various motion planning algorithms. As such, these models are available for public, non-commercial use provided that appropriate reference is made to the model. Also, if you have any mode that you could contribute to this effort, please let us know and we will be happy to post them on this page.

Alpha Puzzle

Provided by Boris Yanovik, GE Corporate Research & Development Center

The alpha puzzle benchmark is a motion planning problem containing a narrow passage. The puzzle consists of two tubes, each twisted into an alpha shape, one tube is the obstacle and the other the moving object (robot). The objective is to separate the intertwined tubes.

Box Folding Problem

Created by Gerson Owen, Darseni I. Gh. Texas A&M University

Possible reasons of a failure...



- Adequate taxonomy for MP problems is complex
 - several dimensions (workspace complexity, C-space dimension, robot typer, etc.)
- What should be compared?
 - Computation time, solution quality (which metrics?)
- Large dependency on implementation details and exploited libraries (collision detection packages)
- Translation using the proprietary file format

from Benchmarking Initiative

... have a look to another community

Automated (Symbolic) Planning



- “my planner is faster than yours”
- difficulty of comparing different planners at the conceptual level
- empirically comparison on a set of available domains: no planner universally dominates all the other planners

... have a look to another community

Automated (Symbolic) Planning



- Adequate taxonomy for planning problems is complex
 - different domains: *deterministic, non-deterministic, probabilistic, temporal, repetitive*
- What should be compared?
 - Correctness, Advice, Performance
- Large dependency on implementation details and exploited libraries
- Several proprietary file format

Their solution: a competition



- Biennial event organized in the context of the International Conference on Planning and Scheduling
- Several goals:
 - analyzing and advancing the state-of-the-art in automated planning systems;
 - providing new data sets to be used by the research community as benchmarks for evaluating different approaches to automated planning;
 - emphasizing new research issues in planning;
 - promoting the acceptance and applicability of planning technology.

Their solution: a competition



- The International Planning Competition has been the driving force in the field since about 1997
- Several systems got an impact on the AI Planning bringing about the key ideas
- URL entry into IPCs: <http://ipc.icaps-conference.org/>

IPC-1 Results



- Not very conclusive.
 - Every planner was good in some respect and not so good in another (nr. problem instances solved, runtime, plan length)
 - The “good” and “bad” depended on both planner and domain
 - The only clear-cut thing was that IPP was faster than SGP in ADL (then again, IPP was in C and SGP in Lisp...)
 - Of course, benchmark results only talk about these benchmarks
 - Still it seemed that generally all these planners were “in the same league”

IPC-2 Main Conclusion



- The planners using heuristics based on “ignoring delete lists”, in particular FF, dramatically outperformed the others.
- The results in previously completely infeasible size regions were considered a breakthrough. The general research focus shifted to the extension to temporal and numeric settings.

IPC-3 Results



- STRIPS versions:
 - “Ignore deletes” planners, particularly FF, still dominant
 - LPG competitive with, and sometimes more efficient than, FF
 - Most of the domains solved quite efficiently
- Temporal non-numeric versions:
 - Efficiency of “ignore deletes” planners, particularly of LPG, can often be retained from STRIPS version
- The left-out numeric variables express the resource constraints!! (like, battery level in Rovers)



IPC-4 Main Conclusion

- Sub-optimal planners yet another step faster through novel heuristics and domain analysis techniques.
- Even relatively realistic domains can often be dealt with relatively well; then again, AS-IS applicability in practice is still far away.
- Optimal planning would be more useful in practice, but optimal planners haven't improved very much over the last 8 years. Seems a lot harder.

Which are the “real” benefits about organizing a competition?



- Three main goals:
 - scientific goal
 - engineering goal
 - community goal

Science, Engineering, and Community
in the Planning Competition
M. Veloso

Scientific Goal



- To serve the advancement of the scientific state of the art of the area.
 - Understand the multiple components of the complete planning problem
 - Impact of different domain representations
 - Complexity of specific goal/state combinations
 - “open” challenge: devising testing problems and evaluation metrics
 - planning in the real world (alpha?)

Engineering Goal



- To contribute to the creation of an actual planning product
- Participants are competitors, but share the goal of advancing the state of the art also in term of good implementations.
- Even if one wins, other competitors excelled at subproblems
 - sharing of code
 - modular implementations

Community Goal



- To assert the research community an integral part of the competition
- Three main aspects:
 - setup of rules
 - promotion of the competition
 - community build up of the competition

Would this also be possible for MP?



- Define a *modeling language*
- Find *benchmarks*
- Define *tracks*
- Define *competition format* and *schedule*
- Define *tools*

Modeling Language



- Automated planning has PDDL (current version 3.0... the language has been advanced faster than the planners)
- Requirement for an input format:
 - **Portable**. Decouple planner data structure and input format object representation.
 - **Flexible**. Support for several robots and obstacles.
 - **Human Readable**. Self explaining and simple to use.
 - **Validable**. Easy to check data consistency and coherence.
 - **Extendible**. Easy to scale to future needs.

Input format for MP



- XML-based tentative input format:
 - Motion Planning Markup Language (MPML) University of Parma
 - Callisto Input Format Utrecht University

- Impact so important that the links are currently not working

Current scenario



- Different formats (.ox, .env, .mif, .iv, .asm, ...), same data:
 - Description of objects either by set of triangles/polyhedra or by union of primitives
 - Description of joints using DH (or a simplified form)
 - Description of kinematic chains as relations between links
- Ever tried to import a benchmark from another tool?
 - New format, new reader (new student?)
 - a pain in the neck

Lack of experimental comparison detrimental to credibility of motion planning research

Modeling Language



- Different formats (.ox, .env, .mif, .iv, .asm, ...), same data:
 - Description of objects either by set of triangles/polyhedra or by union of primitives
 - Description of joints using DH (or a simplified form)
 - Description of kinematic chains as relations between links

Would this also be possible for MP?



- Define a *modeling language*
- **Find *benchmarks***
- Define *tracks*
- Define *competition format* and *schedule*
- Define *tools*

Benchmarks



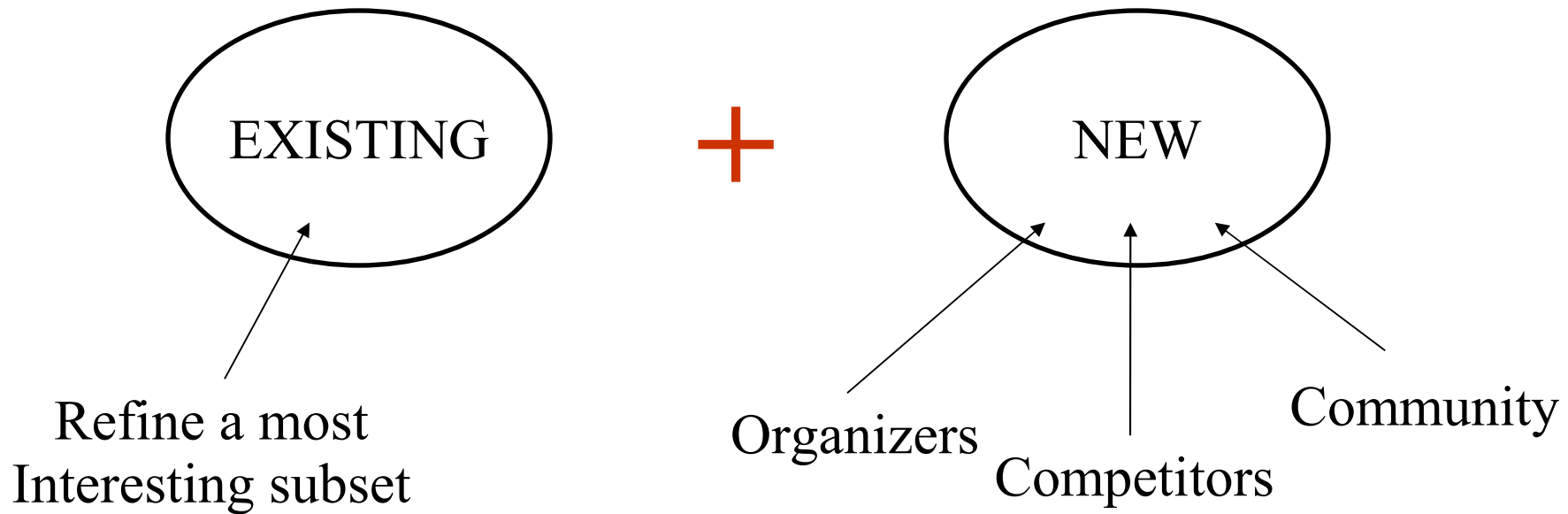
- Keep “old” problems
 - use a wide range of problems
 - reusing many domains from previous competition still hard for current planners
- Traditional route:
 - organizers looking for new problem set
 - organizers need to define:
 - if the domain is “interesting”?
 - significantly “different” from existing benchmarks?

New approach to collect “new” problems



- ask for contribution from competitors:
 - each team submit one domain of their own
 - It's totally ok if that domain is most suitable and performed best by their own planner.
 - So we learn:
 - which type of planner is most suitable for which type of problem
 - and how each planner's best performance compared to others
 - Incentive to submit the domain (more chance to win!)
 - Ask for contribution from community:
 - Open database/wiki for anyone in the planning community to submit a new domain
 - Need a brief description indicating why it's interesting and different from the existing

Benchmarks



Would this also be possible for MP?



- Define a *modeling language*
- Find *benchmarks*
- **Define *tracks***
- Define *competition format* and *schedule*
- Define *tools*

Would this also be possible for MP?



- Define a *modeling language*
- Find *benchmarks*
- Define *tracks*
- Define *competition format and schedule*
- Define *tools*

Competition Format for IPC



Step 1: Organizers provide Language Manual

Step 2: Organizers provide Test Problems

- Not belongs to benchmarks used to grade the planners
- Same level of expressiveness

Step 3: Competitors Submit Planners

Step 4: Organizers run planners and collect/analyze results

- Testing problems are not known to the competitors (blind evaluation)



After the Competition

- Used problem sets will be available
- Results/Analysis
- Planner descriptions
- Also strongly suggest the competitor to contribute the planner code
 - facilitate more collaborations

Would this also be possible for MP?



- Define a *modeling language*
- Find *benchmarks*
- Define *tracks*
- Define *competition format* and *schedule*
- **Provide *tools***

Tools



- C/C++ Java Parser for input format
- Collision detection libraries
- Sampling techniques

Sources



- Repository in the form of Wiki (Mark Roberts's talk at 16:50):
 - New announcement from organizer
 - Schedule change
 - Language manual
 - Test problems
 - Results
 - Vision: will eventually extended to have anything and everything that planning researchers care about:
 - Update on benchmark problems
 - Results: complexity, new improved solution quality
 - Description of different planning topics and References to related work
 - etc.

Schedule



- Final modelling language scope: Dec 2007
- Test problems: Jan - Feb
- Finalize problem set: Mar - April
- Collecting Results: May - June
 - Expect lot of problems and runs
- Result analysis: July - Aug
- Note: optimistic schedule that can be shifted up to a month later

Lot of work: is there a community?



- IEEE RAS Technical Committee (88 members)

Algorithms for Planning and Control of Robot Motion

Main :: Home Page

Edit History Print

- ◆ Home
- ◆ News
- ◆ Events
- ◆ Literature
- ◆ PhD Theses
- ◆ Success Stories
- ◆ Video Showcase

IEEE RAS Technical Committee on Algorithms for Planning and Control of Robot Motion

Scope
As modern robots address real-world problems in dynamic, unstructured, and open environments, novel challenges arise in the areas of robot control algorithms and motion planning. These challenges stem from an increased need for autonomy and flexibility in robot motion and task

- So what are we waiting for?

Do we need a competition in MP?



- Competitions are actually “needed”?
- “winning” a competition at some time could be the result of factors that may not be the real long-term relevant factors
- many advances in science have been achieved without competitions :)
- They are:
 - an aspect of the current days
 - they have been relevant to advance the state of the art of several fields