RATIONAL DEPLOYMENT OF CSP HEURISTICS

IJCAI 2011

David Tolpin, Solomon Eyal Shimony
Ben Gurion University of the Negev,
Beer Sheva, Israel

Background

Constraint Satisfaction

A constraint satisfaction problem (CSP) is defined by:

variables $X = \{X_1, X_2, \dots\}$,
constraints $C = \{C_1, C_2, \dots\}$

- Each *variable* X_i has a non-empty domain D_i of possible values.
- Each *constraint* C_i involves some subset of the variables—the *scope* of the constraint—and specifies the allowable combinations of values for that subset.
- An *assignment* that does not violate any constraints is called *consistent* (or solution).

Rational Metareasoning

- A problem-solving agent can perform *base-level* actions from a known set $\{A_i\}$.
- Before committing to an action, the agent may perform a sequence of *meta-level* deliberation actions from a set $\{S_j\}$.
- At any given time there is a base-level action A_α that maximizes the agent's *expected utility*.

The **net VOI** $V(S_j)$ of action S_j is the intrinsic VOI Λ_j less the cost of S_j:

$$V(S_j) = \Lambda(S_j) - C(S_j)$$

The **intrinsic VOI** $\Lambda(S_j)$ is the expected difference between the intrinsic expected utilities of the new and the old selected base-level action, computed after the meta-level action is taken:

$$\Lambda(S_j) = \mathrm{E}[\mathrm{E}U(A_\alpha^j) - \mathrm{E}U(A_\alpha)]$$

- $S_{j_{\max}}$ that maximizes the net VOI is performed: $j_{\max} = \arg\max_j V(S_j)$ if $V(S_{j_{\max}}) > 0$.
- Otherwise, A_α is performed.

Overview

A heuristic must be both informative and efficient to compute. Overhead of some well-known heuristics may outweigh the gain. Such heuristics should be deployed adaptively.

Case Study

- CSP backtracking search algorithms typically employ variable-ordering and value-ordering heuristics.
- Some value ordering heuristics are computationally heavy, e.g. heuristics based on solution count estimates.
- Principles of rational metareasoning can be applied to decide when to deploy the heuristics.

Value Ordering

Value ordering heuristics convey information about:

- T_i—the expected time to find a solution with $X_k = y_{ki}$;
- p_i—the probability that there is no solution with $X_k = y_{ki}$.

The expected remaining search time in the subtree under X_k for ordering ω is:

$$T^{s|\omega} = T_{\omega(1)} + \sum_{i=2}^{|D_k|} T_{\omega(i)} \prod_{j=1}^{i-1} p_{\omega(j)}$$

- The current optimal base-level action is picking the ω which minimizes $T^{s|\omega}$.
- The intrinsic VOI Λ_i of estimating T_i, p_i for the ith assignment is the expected decrease in $T^{s|\omega}$: $\Lambda_i = \mathrm{E}[T^{s|\omega_-} - T^{s|\omega_{+i}}]$.
- Computing new estimates (with overhead T^c) for values T_i, p_i is beneficial when the net VOI is positive: $V_i = \Lambda_i - T^c$.

Main Results

Rational Value Ordering

The intrinsic VOI Λ_i of invoking the heuristic can be approximated as:

$$\Lambda_i \approx \mathrm{E}[(T_1 - T_i)|D_k| \,\big|\, T_i < T_1]$$

VOI of Solution Count Estimates

The net VOI V of estimating a solution count can be approximated as:

$$V \propto |D_k| e^{-\nu} \sum_{n=n_{\max}}^{\infty} \left(\frac{1}{n_{\max}} - \frac{1}{n}\right) \frac{\nu^n}{n!} - \gamma$$

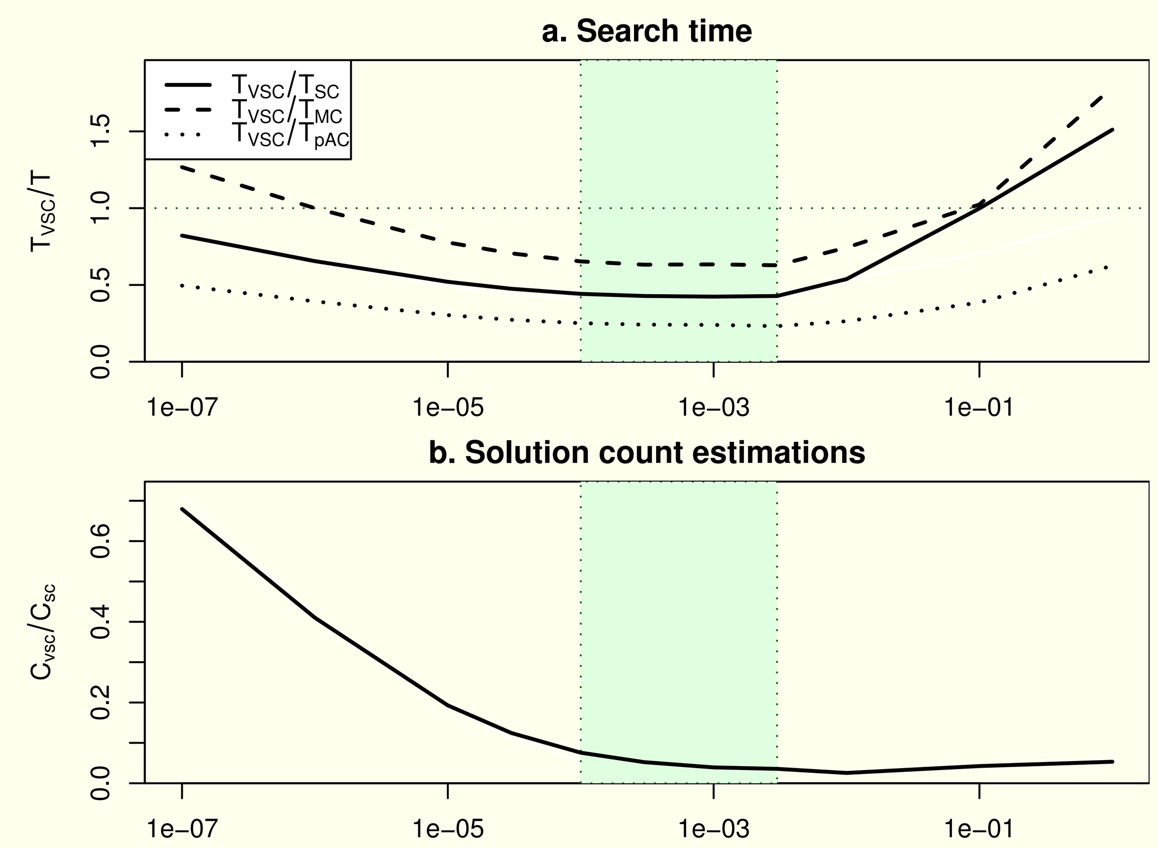
where the constant γ depends on the search algorithm and the heuristic, rather than on the CSP instance, and can be learned offline.

Experiments

Benchmarks

14 benchmarks from CSP Solver Competition 2005:

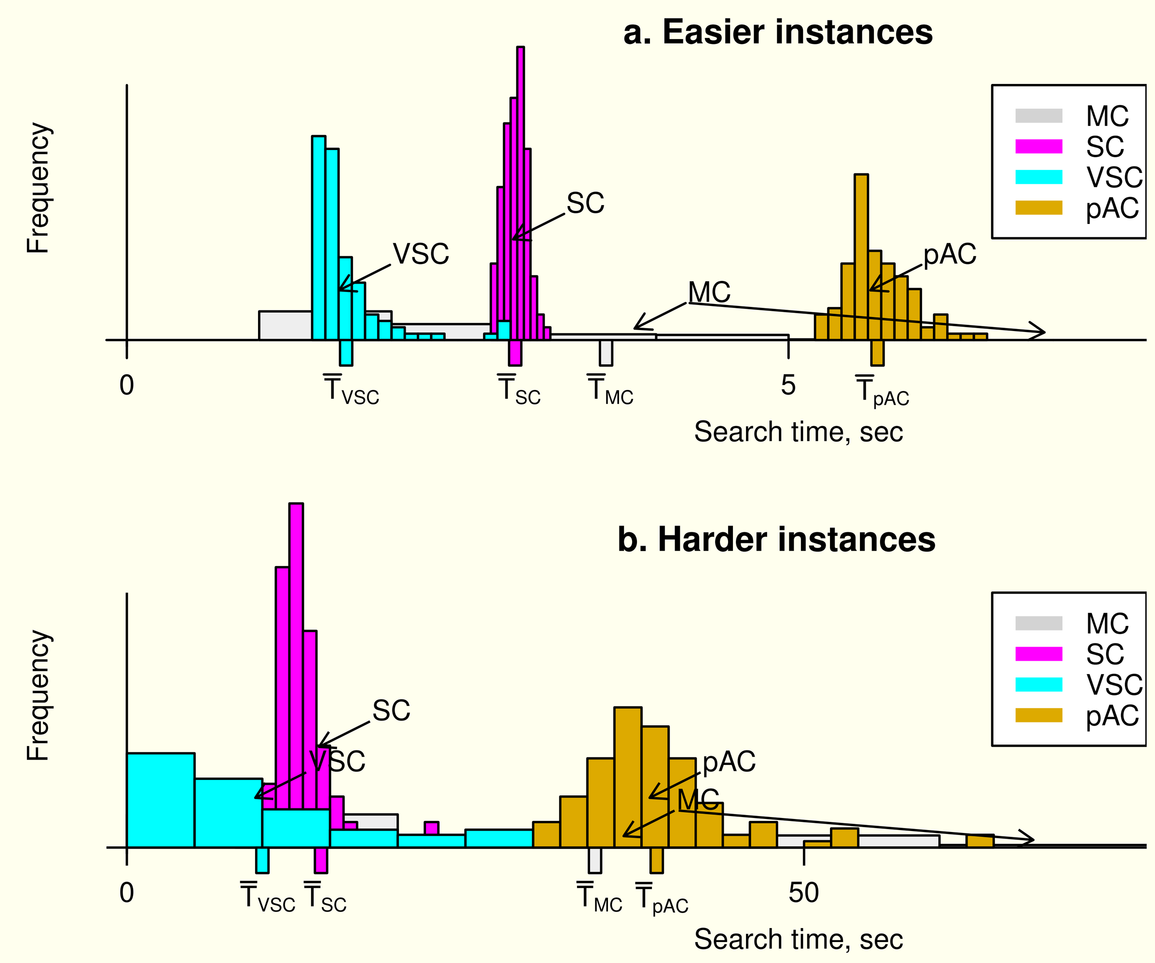
- for $\gamma = 0$;
- for the range $\gamma \in \{10^{-7}, 10^{-6}, \dots, 1\}$,
- with the *minimum-conflicts* and the *pAC* heuristics.



The maximum improvement is achieved when the solution count is estimated in a small fraction of occasions.

Random instances (Model RB)

Exhaustive deployment, rational deployment, the *minimum conflicts* and the *pAC* heuristics were compared on two sets of 100 problem instances.



$\gamma = 10^{-3}$ based on a small set of hard instances gave good results on sets of instances of varying size and hardness.

Generalized Sudoku

- Real-world problem instances have much more structure than Model RB random instances.
- The experiments were repeated on random Generalized Sudoku instances — a highly structured domain.
- Relative performance was similar to Model RB.

Summary

- A model for adaptive deployment of value ordering heuristics in algorithms for constraint satisfaction problems.
- Steady improvement compared to exhaustive deployment for an heuristic based on solution count estimates.
- The optimum performance is achieved when solution counts are estimated only in a small number of search states.

Future Work

- Generalization of the VOI to deploy different types of heuristics for CSP.
- Explicit evaluation of the quality of the distribution model, coupled with a better candidate model of the distribution.
- Application to search in other domains, especially to heuristics for planning; in particular, examining whether the meta-reasoning scheme can improve reasoning over deployment based solely on learning.

Acknowledgments

- IMG4 Consortium under the MAGNET program of the Israeli Ministry of Trade and Industry
- Lynne and William Frankel Center for Computer Sciences
- Israel Science Foundation grant 305/09
- Paul Ivanier Center for Robotics Research and Production Management