RATIONAL DEPLOYMENT OF CSP HEURISTICS

IJCAI 2011

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

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:

$$\Lambda(S_j) = \mathrm{E}(\mathrm{E}(U(A^j_\alpha)) - \mathrm{E}(U(A_\alpha)))$$

- $S_{j_{\max}}$ 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.

Acknowledgments

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

Overview

Heuristics are crucial tools in decreasing search effort in various areas of AI.
A heuristic must provide useful information to the search algorithm, and be efficient to compute.
Overhead of some well-known heuristics may outweigh the gain.
Such heuristics should be deployed selectively, based on principles of rational metareasoning.

Case Study

- CSP backtracking search algorithms typically employ variable-ordering and value-ordering heuristics.
- Many 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 provide information about:

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

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

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

- The current optimal base-level action is picking the ω which optimizes $T^{s|\omega}$. $T^{s|\omega}$ is minimal if the values are sorted by increasing order of $\frac{T_i}{1-p_i}$.
- The intrinsic VOI Λ_i of estimating T_i, p_i for the ith assignment is the expected decrease in the expected search time: $\Lambda_i = \mathrm{E}[T^{s|\omega_-} - T^{s|\omega_{+i}}]$.
- Computing new estimates (with overhead T^c) for values T_i, p_i is beneficial just 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:

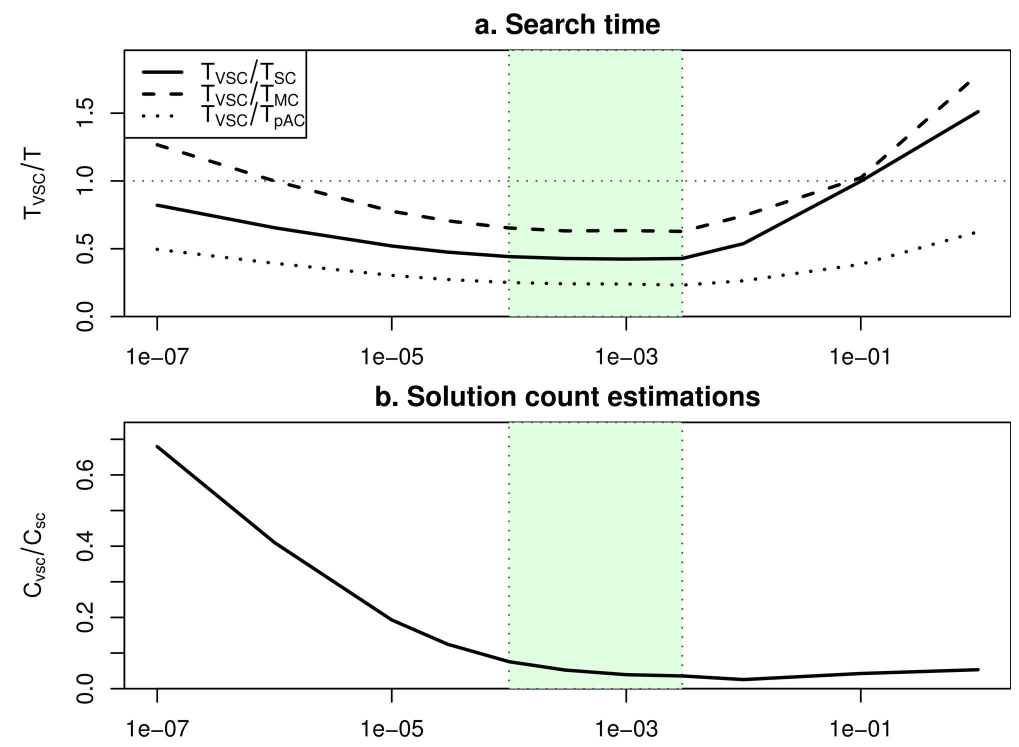
$$V \propto |D_k| e^{-\nu} \sum_{n=n_{\max}}^{\infty} \left(\frac{1}{n_{\max}} - \frac{1}{n}\right) \frac{\nu^n}{n!} - \gamma$$

where the constant γ depends on the search algorithm and the heuristic, rather than on the CSP instance, and can be learned offline. The infinite sum $\sum_{n=n_{\max}}^{\infty} \left(\frac{1}{n_{\max}} - \frac{1}{n}\right) \frac{\nu^n}{n!}$ is rapidly converging and can be computed efficiently.

Experiments

Benchmarks

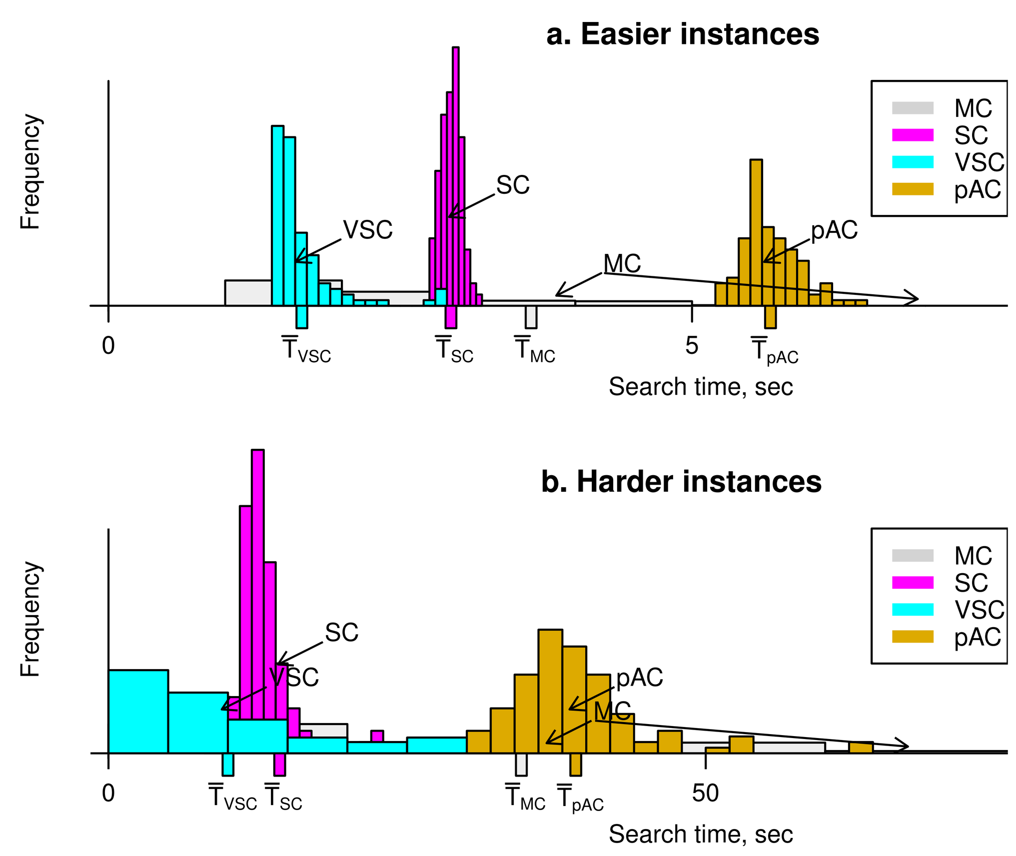
CSP benchmarks from CSP Solver Competition 2005 were used. 14 benchmarks were solved for $\gamma = 0$ and the exponential range $\gamma \in \{10^{-7}, 10^{-6}, \dots, 1\}$, as well as with the minimum-conflicts heuristic and the pAC heuristic.



The maximum improvement is achieved when the solution count is estimated only in a small fraction of occasions selected using rational metareasoning.

Random instances

Based on the results on benchmarks, we chose $\gamma = 10^{-3}$, and applied it to two sets of 100 problem instances. Exhaustive deployment, rational deployment, the minimum conflicts heuristic, and the pAC heuristic were compared.



The value of γ chosen based on a small set of hard instances gave good results on a set of instances with different parameters and of varying hardness.

Generalized Sudoku

- Real-world problem instances often have much more structure than random instances generated according to Model RB.
- We repeated the experiments on randomly generated Generalized Sudoku instances— a highly structured domain.
- Relative performance on Generalized Sudoku was similar to Model RB.

Conclusions

- This work suggests a model for adaptive deployment of value ordering heuristics in algorithms for constraint satisfaction problems.
- As a case study, the model was applied to a value-ordering heuristic based on solution count estimates, and a steady improvement was achieved compared to always computing the estimates.
- For many problem instances 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.