Making Good Decisions Quickly

Sven Koenig, *Senior Member, IEEE*

Abstract—Several disciplines, including artificial intelligence, operations research and many others, study how to make good decisions. In this overview article, we argue that the key to making progress in our research area is to combine their ideas, which often requires serious technical advances to reconcile their different assumptions and methods in a way that results in synergy among them. To illustrate this point, we give a broad overview of our ongoing research on search and planning (with a large number of students and colleagues, both at the University of Southern California and elsewhere) to demonstrate how to combine ideas from different decision making disciplines. For example, we describe how to combine ideas from artificial intelligence, operations research, and utility theory to create the foundations for building decision support systems that fit the risk preferences of human decision makers in high-stake one-shot decision situations better than current systems. We also describe how to combine ideas from artificial intelligence, economics, theoretical computer science and operations research to build teams of robots that use auctions to distribute tasks autonomously among themselves, and give several more examples.

Index Terms—agents, ant robotics, artificial intelligence, auction-based coordination, decision theory, dynamic programming, economics, freespace assumption, goal-directed navigation, greedy online planning, heuristic search, high-stake one-shot decision making, incremental heuristic search, Markov decision processes, multi-agent systems, nonlinear utility functions, operations research, planning, real-time heuristic search, reinforcement learning, risk preferences, robotics, scarce resources, sequential-single item auctions, terrain coverage, utility theory.

I. INTRODUCTION

ARTIFICIAL INTELLIGENCE is rooted in building cognitive systems (that is, systems that operate in a way similar to the human mind) but today is more and more about engineering intelligent systems (that is, systems that solve tasks that require difficult decisions) even if these systems do not operate in a way similar to the human mind. For example, the popular textbook "Artificial Intelligence: A Modern Approach" [1] by Stuart Russell and Peter Norvig views artificial intelligence as the science of creating rational agents, where agents are control systems that interact with an environment. They can sense to gather information about the state of the environment and execute actions to change it. Rational agents, according to the textbook, should select actions that are expected to maximize given performance measures. In general, agents must be able to make good decisions in complex situations that involve a substantial degree of uncertainty, yet find solutions in a timely manner. Researchers from artificial intelligence therefore create a strong foundation for building such agents, typically focusing more on autonomous decision making and optimization than modeling of complex decision

Sven Koenig is with the Department of Computer Science, University of Southern California, Los Angeles, CA, 90089-0781, USA, e-mail: skoenig@usc.edu, web page: idm-lab.org.

problems or providing decision support for human decision makers.

Artificial intelligence has developed tools for building agents that perform well with respect to given performance measures. Other decision making disciplines provide different and potentially complementary tools. In general, the larger one's toolbox, the more decision problems one is able to tackle. By combining ideas from different decision making disciplines, one can expect to improve on existing tools and build new tools that either perform better than existing ones or solve decision problems that existing tools cannot solve. This provides an incentive to study different decision making disciplines, develop curricula that allow students to learn about several decision making disciplines, and create a universal science of intelligent decision making that combines ideas from different decision making disciplines, including artificial intelligence, operations research, economics, decision theory, and control theory. One obstacle that needs to be overcome is that different decision making disciplines typically study different applications and thus make different assumptions, resulting in different decision making methods. Combining ideas from different decision making disciplines therefore often requires serious technical advances to reconcile the different assumptions and methods in a way that results in synergy among them. A second obstacle is that that different decision making disciplines focus on different aspects of decision problems and have different ideas about what constitutes a good solution to a given decision problem, often due to the disciplinary training of their researchers. For example, statistics researchers often tend to focus on the uncertainty in the data and how it can be resolved; optimization researchers often tend to assume that the data is correct and focus on finding optimal or close to optimal (rather than timely) solutions (concentrating on "planning" rather than "operations"); and artificial intelligence researchers often tend to focus on the ability of agents to make good decisions online, taking into account the limitations of the agents (such as their limited sensing, computational and communication capabilities as well as their noisy actuation) in addition to their interaction with the environment (such as information collection) and each other (such as coordination), which explains the title of this overview article. A third obstacle is that different decision making disciplines often use different terminology and notation. Multi-disciplinary training can overcome these obstacles and transform the second obstacle into a strength.

Artificial intelligence often pursues general principles that apply widely to decision making and problem solving (rather than problem-specific methods), perhaps due to its roots in building cognitive systems. It is therefore not surprising that artificial intelligence, over time, has incorporated ideas from other decision making disciplines. For example, the third

edition of "Artificial Intelligence: A Modern Approach" covers local search in Chapter 4, including hill-climbing search, simulated annealing, local beam search, and genetic algorithms. It covers utility theory in Chapter 16, including utility functions, multi-attribute utility functions, and influence diagrams. It covers sequential decision problems in Chapter 17, including Markov decision processes and dynamic programming methods such as value and policy iteration. It covers game theory in Chapter 17, including single-move, repeated, and sequential games. It also covers mechanism design in the same chapter, including auctions. All of these topics have also been studied in other decision making disciplines, such as operations research and economics, and typically originated there. For example, researchers from artificial intelligence discovered totally and partially observable Markov decision processes from operations research when working on foundations for decision theoretic planning and reinforcement learning and then, for example, developed new ways of representing and solving them by incorporating insights from knowledge representation and planning (where states are typically represented as collections of facts), resulting in both symbolic and structured dynamic programming. Symbolic dynamic programming, for example, is a generalization of dynamic programming for solving Markov decision processes that exploits symbolic structure in the solution of relational and first-order logical Markov decision processes to avoid the full state and action enumeration of classical dynamic programming methods [2].

Outsiders often do not know about these and other recent achievements of artificial intelligence and, for this reason, might not appreciate the ideas that it has to offer to them. There exist some established but narrow interfaces between artificial intelligence and other decision making disciplines. An example of a step in the direction of an interface between artificial intelligence and control theory is [3]. An example of a step in the direction of an interface between artificial intelligence and operations research is the International Conference on Integration of Artificial Intelligence and Operations Research Techniques in Constraint Programming for Combinatorial Optimization Problems (CPAIOR), which is by now an established conference series with 9 conferences since 2004, preceeded by 5 workshops. Similarly, ILOG eventually integrated software for constraint programming and linear optimization. In general, however, artificial intelligence probably needs to reach out even more to other decision making disciplines with the objective to inform them and create a universal science of intelligent decision making. While this might appear to be an obvious objective, progress in this direction has been made mostly recently. For example, an algorithmic decision theory community formed around 2000 and eventually created the International Conference on Algorithmic Decision Theory (ADT). The First International Conference on Algorithmic Decision Theory took place in Venice, Italy, in 2009, and the Second International Conference on Algorithmic Decision Theory took place in New Brunswick, USA, in 2011. The conference series, according to the conference announcement at www.adt2011.org, involves researchers from such disparate fields as decision theory, discrete mathematics, theoretical computer science, economics,

and artificial intelligence, aiming to improve decision support in the presence of massive data bases, partial and/or uncertain information, and distributed decision makers. Papers have covered topics from computational social choice to preference modeling, from uncertainty to preference learning, and from multi-criteria decision making to game theory [4].

We sketch some of our own research in the remainder of this overview article to illustrate why we believe that it is important to combine ideas from different decision making disciplines. Not surprisingly, our research centers around methods for decision making (planning and learning) that enable single agents and teams of agents to act intelligently in their environments and exhibit goal-directed behavior in real-time, even if they have only incomplete knowledge of their environment, imperfect abilities to manipulate it, limited or noisy perception or insufficient reasoning speed. Our research group, the Intelligent Decision Making group, develops new decision making methods, implements them and studies their properties theoretically and experimentally. We demonstrated around 1995 that it is possible to combine ideas from different decision making disciplines by developing a robot navigation architecture based on partially observable Markov decision processes from operations research that allows robots to navigate robustly despite a substantial amount of actuator and sensor uncertainty, which prevents them from knowing their precise location during navigation [5]. This research resulted in a reliable robot architecture that overcomes the deficiencies of purely topological or metric navigation methods [6]. Since then, our research group has continued to combine ideas from different decision making disciplines. In the following, we describe some of these research directions in more detail. While they might appear diverse, there is a common underlying thrust, namely to bring about advances that extend the reach of search (in a broad sense, including heuristic search, hill-climbing and dynamic programming), and to apply the results to robot navigation.

II. EXAMPLE: NONLINEAR UTILITY FUNCTIONS

Finding plans that maximize the expected utility for nonlinear utility functions is important in both high-stake one-shot decision situations and decision situations with scarce resources [7].

- In high-stake one-shot decision situations, huge gains or losses of money or equipment are possible, and human decision makers take risk aspects into account. Risk-averse decision makers, for example, tolerate a smaller expected plan-execution reward for a reduced variance (although this explanation is a bit simplified). For example, they try to avoid huge losses when fighting forest fires, containing marine oil spills or controlling autonomous spacecraft (and other decision problems that artificial intelligence researchers study) and thus add more sensing operations than necessary to maximize the expected reward [8]. Planning systems need to reflect these risk preferences. Bernoulli and Von Neumann/Morgenstern's utility theory [9] [10] suggests that rational human decision makers choose plans that

maximize the expected utility, where the utility is a monotonically increasing function of the reward. For example, exponential utility functions completely preserve the structure of planning tasks because they are the only class of nonlinear utility functions for which decisions do not depend on the accumulated reward. However, one-switch utility functions often model the risk attitudes of human decision makers better than exponential utility functions [11].

- In decision situations with scarce resources, there are often limits to how much of a resource (such as time, energy or memory) can be consumed before it runs out. For example, a lunar rover that reaches a science target with minimal expected energy consumption does not necessarily maximize the probability of achieving it within its battery limit. Resource limits can be modeled with monotonically (but perhaps not strictly monotonically) increasing utility functions that map total rewards (the negative of the total resource consumptions) to real values. For example, a hard resource limit can be modeled with a step function that is zero to the left of the negative resource limit (where the total resource consumption is greater than the resource limit) and one to the right of it [12].

Decision-theoretic planning methods in artificial intelligence are these days typically, either explicitly or implicitly, based on Markov decision processes. One can use dynamic programming methods, such as value iteration [13] or policy iteration [14], to maximize the expected total (undiscounted or discounted) reward. One can also use these methods to maximize the expected utility for nonlinear utility functions (studied in the context of risk-sensitive Markov decision processes in operations research [15] and control theory [16]) but then, except for exponential utility functions, needs to add the accumulated reward to the states, which increases the number of states substantially. We and other artificial intelligence researchers have therefore studied "functional" versions of value and policy iteration that do not maintain a value for each augmented state but rather a value function for each original state (that maps the total reward to the value of the state) and operate directly on these value functions [17] [18] [19] [20] [21], which allows one to solve larger decision problems than what would be possible otherwise due to the following advantages: First, the value functions can sometimes be represented exactly and compactly (that is, with a finite number of parameters), as we have shown for one-switch utility functions [22] [23] and piecewise linear utility functions with optional exponential tails [24]. Second, the value functions can also be approximated to a desired degree (for example, with piecewise linear functions), sometimes resulting in approximation guarantees, which allows one to trade off between runtime and memory consumption on one hand and solution quality on the other hand [24]. More complex decision problems can be solved in a similar way. For example, a lunar rover might have to maximize its science return within its battery limit despite uncertainty about its energy consumption, when scientists have designated several locations that the rover can visit to perform science experiments and assigned a science return value to each of them [25]. Other approaches also exist [26], together with extensions to teams of robots [27].

Methods from artificial intelligence exploit the structure of decision-theoretic planning tasks [28]. For example, artificial intelligence has investigated how to represent search spaces implicitly and exploit the resulting decomposability to solve Markov decision processes efficiently without having to enumerate their state spaces completely. For instance, structured versions of value iteration represent the transition policies in factored form, which allows them to represent policies more compactly than with tables to speed up their computations and generalize policies across states [29]. An example is SPUDD, that uses algebraic decision diagrams instead of tables [30]. Artificial intelligence has also investigated forward search methods that, different from value and policy iteration, consider only states that are reachable from the start state. For instance, LAO* uses heuristic search to restrict the value updates only to relevant states rather than all states [31] [32]. We have generalized these methods to find plans that maximize the expected utility for nonlinear utility functions [33]. Other decision making disciplines have developed other ways of exploiting the structure of decision-theoretic planning tasks [34], meaning that there are opportunities for combining different ideas. Overall, this research combines insights from artificial intelligence, operations research, and utility theory for planning with nonlinear utility functions. Operations research has studied the properties of Markov decision processes in detail, artificial intelligence and operations research contribute ideas for solving them, and utility theory provides a realistic optimization criterion for high-stake one-shot decision situations.

III. EXAMPLE: AUCTION-BASED COORDINATION

Centralized control is often inefficient for teams of robots in terms of the amount of communication and computation required since the central controller is the bottleneck of the system. Researchers from artificial intelligence and robotics have therefore studied robot coordination with cooperative auctions [35]. An auction is "a market institution with an explicit set of rules determining resource allocation and prices on the basis of bids from the market participants" [36]. Auctions have been developed for the allocation of resources in situations where agents have different utilities and private information. Auctions are therefore promising decentralized methods for teams of robots to allocate and re-allocate tasks in real-time among themselves in dynamic, partially known and time-constrained domains with positive or negative synergies among tasks. Furthermore, the short length of a bid is helpful when communication bandwidth is limited. Artificial intelligence and later robotics have explored auction-based coordination systems at least since the introduction of contract networks [37], mostly from an experimental perspective. In auction-based coordination systems, the bidders are robots, and the items up for auction are tasks to be executed by the robots. All robots bid their costs. Thus, the robot with the smallest bid

cost is best suited for a task. All robots then execute the tasks that they win. Auction-based coordination systems are easy to understand, simple to implement and broadly applicable. They promise to be efficient both in communication (since robots communicate only essential summary information) and in computation (since robots compute their bids in parallel). A typical application is multi-robot routing [38], where a team of robots has to visit given targets and repeatedly reassigns targets among the robots as it learns more about the initially unknown terrain, as robots fail or as additional targets get introduced. Examples include environmental clean up, mine clearing, space exploration, and search-and-rescue. Multi-robot routing problems are NP-hard to solve optimally even if the locations of obstacles, targets, and robots are initially known and (except for the locations of the robots) do not change [39]. Their similarity to traveling salesperson problems [40] allows one to use insights from theoretical computer science and operations research for their analysis. Economics has an extensive auction literature but its agents are rational and competitive, leading to long decision cycles, strategic behavior, and possibly collusion. Such issues do not arise in auction-based coordination systems because the robots faithfully execute their programs. On the other hand, auction-based coordination systems must operate in real-time. Still, some insights from economics can be exploited for building them, such as the concepts of synergy and different auction mechanisms, including parallel, combinatorial, and sequential single-item (SSI) auctions. For example, SSI auctions proceed in several rounds, assigning one additional target per round to some robot. We have exploited the fact that SSI auction-based coordination systems with marginal-cost bidding [41] perform a form of hill-climbing search to analyze the resulting team performance [42]. We have used tools from theoretical computer science to show that SSI auction-based coordination systems can provide constant factor performance guarantees even though they run in polynomial time and, more generally, that they combine advantageous properties of parallel and combinatorial auctions [39], resulting in one of the few existing performance analyses. Some intuition for this result can be gained from interpreting the greedy construction of minimum spanning trees as a cooperative auction [43]. We have investigated several versions of SSI auctions to build SSI auction-based coordination systems that increase the team performance while still allocating targets to robots in real-time. For example, we have generalized auction-based coordination systems based on SSI auctions to assign more than one additional target during each round (called the bundle size), which increases their similarity with combinatorial auctions by taking more synergies among targets into account and making the resulting hill-climbing search less myopic. We have shown that, for a given number of additional targets to be assigned during each round, every robot needs to submit only a constant number of bids per round and the runtime of winner determination is linear in the number of robots [44]. Thus, the communication and winner determination times do not depend on the number of targets, which helps the resulting auction-based coordination systems to scale up to a large number of targets for small bundle sizes. Overall, this research combines insights from artificial intelligence, economics, theoretical computer science and operations research for the development of auction-based coordination systems and their analysis [45].

IV. EXAMPLE: FAST REPLANNING

Robots often operate in domains that are only incompletely known or change over time. One way of dealing with incomplete information is to interleave search with action execution. In this case, the robots need to replan repeatedly. To make search fast, one can use heuristic search methods with limited lookahead (agent-centered search, such as real-time heuristic search [46]) or heuristic search methods that reuse information from previous searches (incremental heuristic search). Consider, for example, a robot that has to move from its current location to given goal coordinates in initially unknown terrain. The robot does not know the locations of obstacles initially but observes them within its sensor radius and adds them to its map. Planning in such non-deterministic domains is typically time-consuming due to the large number of contingencies, which provides incentive to speed up planning by sacrificing the optimality of the resulting plans. Greedy online planning methods interleave planning and plan execution to allow robots to gather information early and then use the acquired information right away for replanning, which reduces the amount of planning performed for unencountered situations. For example, goal-directed navigation with the freespace assumption is a common-sense version of assumption-based planning that is popular in robotics for moving a robot to a given goal location in initially unknown terrain [47] and can be analyzed with tools from theoretical computer science [48]. It finds a short (unblocked) path from the current location of the robot to the goal location given its current knowledge of the locations of obstacles under the assumption that the terrain is otherwise free of obstacles. If such a path does not exist, it stops unsuccessfully. Otherwise, the robot follows the path until it either reaches the goal location, in which case it stops successfully, or observes the path to be blocked, in which case it repeats the process using its revised knowledge of the locations of obstacles. Incremental heuristic search methods solve such series of similar path planning problems often faster than searches from scratch [49] (by reusing information from previous searches to speed up their current search), yet differ from other replanning methods (such as planning by analogy) in that their solution quality is as good as the solution quality of searches from scratch [50]. The first incremental heuristic search methods was published in artificial intelligence and robotics [51]. It has been discovered since then that incremental search had been studied much earlier already (for example, in the context of dynamic shortest path problems in algorithms), which allowed us to develop a new incremental heuristic search method by combining ideas from different disciplines. D* Lite [52] is now a popular incremental heuristic search method for planning with the freespace assumption that combines ideas from incremental search (namely, to recalculate only those start distances that can have changed or have not been calculated before) with ideas from heuristic search (namely, to use approximations of the goal distances

to recalculate only those start distances that are relevant for recalculating a shortest path). In particular, it combines ideas behind DynamicSWSF-FP [53] from algorithms with ideas behind A* from artificial intelligence. Overall, this research combines insights from artificial intelligence, robotics, and theoretical computer science for the development of fast replanning methods and their analysis.

V. EXAMPLE: ANT ROBOTS

Researchers from robotics are interested in simple robots with limited sensing and computational capabilities as well as noisy actuation. Such ant robots have the advantage that they are easy to program and cheap to build. This makes it feasible to deploy groups of ant robots and take advantage of the resulting fault tolerance and parallelism. Researchers from robotics had studied robots that can follow trails laid by other robots but we studied robots that leave trails in the terrain to cover closed terrain (that is, visit each location) once or repeatedly, as required for surveillance, guarding terrain, mine sweeping, and surface inspection. Ant robots cannot use conventional planning methods due to their limited sensing and computational capabilities. To overcome these limitations, we developed navigation methods that leave markings in the terrain, similar to the pheromone trails of real ants. These markings are shared among all ant robots and allow them to cover terrain even if they do not have any kind of memory, cannot maintain maps of the terrain, nor plan complete paths. They can be used by single ant robots as well as groups of ant robots and provide robustness in situations where some ant robots fail, ant robots are moved without realizing this, the trails are of uneven quality, and some trails are destroyed. Robot architectures based on partially observable Markov decision processes provide robots with the best possible location estimate to overcome actuator and sensor uncertainty, while ant robots achieve their goals without ever worrying about where they are in the terrain. We built physical ant robots that cover terrain and test their design both in realistic simulation environments and on a Pebbles III robot. We modeled the coverage strategy of such ant robots with graph dynamic programming methods that are similar to real-time heuristic search methods (such as Learning Real-Time A*) [46] and reinforcement learning methods (such as Real-Time Dynamic Programming) [54] from artificial intelligence (except that the values are written on the floor rather than stored in memory), which allowed us to use tools from theoretical computer science to analyze their behavior [55]. Other researchers, such as Israel Wagner and his collaborators, have similar interests and work on the intersection of robotics, artificial intelligence, and theoretical computer science [56], see also http://www.cs.technion.ac.il/~wagner/. Overall, this research combines insights from artificial intelligence, robotics, biology, and theoretical computer science for the development of navigation methods for ant robots and their analysis.

VI. EXAMPLE: TERRAIN COVERAGE

Robot coverage of known terrain can be sped up with multiple robots that coordinate explicitly. Researchers from robotics had investigated spanning tree-based coverage methods in unweighted terrain, where the travel times of robots are the same everywhere in the terrain. Single-robot coverage problems are solved with minimal cover times by Spanning Tree Coverage (STC), a polynomial-time single-robot coverage method published in robotics and artificial intelligence that decomposes terrain into cells, finds a spanning tree of the resulting graph, and makes the robot circumnavigate it [57] [58]. This method had been generalized to Multi-Robot Spanning Tree Coverage (MSTC), a polynomial-time multi-robot coverage method published in robotics [59] [60]. While MSTC provably improves the cover times compared to STC, it cannot guarantee its cover times to be small. We showed that solving several versions of multi-robot coverage problems with minimal cover times is NP-hard, which provides motivation for designing polynomial-time constant-factor approximation methods. We generalized STC to Multi-Robot Forest Coverage (MFC), a polynomial multi-robot coverage method based on a method published in operations research [61] (in the context of deciding where to place nurse stations in hospitals) for finding tree covers with trees of balanced weights, one tree for each robot. We also generalized MFC from unweighted terrain to weighted terrain, where the travel times of robots are not the same everywhere. The cover times of MFC in weighted and unweighted terrain are at most about sixteen times larger than minimal and experimentally close to minimal in all tested scenarios [62]. Overall, this research combines insights from artificial intelligence, robotics, and operations research for the development of terrain coverage methods and their analysis.

VII. CONCLUSIONS

In this overview article, we described some of our own research to illustrate why we believe that it is important to combine ideas from different decision making disciplines. We are convinced that we have overlooked lots of developments but encourage researchers from artificial intelligence to continue to reach out to other decision making disciplines with the objective to inform them about our latest research and help to make progress towards a universal science of intelligent decision making.

ACKNOWLEDGMENTS

REFERENCES

[1] S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*, 3rd ed. Prentice Hall, 2009.

[2] S. Sanner and K. Kersting, "Symbolic dynamic programming," in *Encyclopedia on Machine Learning*. Springer, 2010, pp. 946–954.

[3] T. Dean, *Planning and Control*. M. Kaufmann Publishers, 1991.

[4] F. Rossi and A. Tsoukias, Eds., *Algorithmic Decision Theory, First International Conference, ADT 2009*, ser. LNAI 5783. Springer, 2009.

[5] S. Koenig and R. Simmons, "Xavier: A robot navigation architecture based on partially observable markov decision process models," in *Artificial Intelligence Based Mobile Robotics: Case Studies of Successful Robot Systems*, D. Kortenkamp, R. Bonasso, and R. Murphy, Eds. MIT Press, 1998, pp. 91–122.

[6] S. Thrun, W. Burgard, and D. Fox, Eds., *Probabilistic Robotics*. MIT Press, 2005.

[7] J. Bresina, R. Dearden, N. Meuleau, S. Ramakrishnan, D. Smith, and R. Washington, "Planning under continuous time and resource uncertainty: A challenge for AI," in *Proceedings of the Conference on Uncertainty in Artificial Intelligence*, 2002, pp. 77–84.

[8] S. Koenig and Y. Liu, "Sensor planning with non-linear utility functions," in *Proceedings of the European Conference on Planning*, 1999, pp. 265–277.

[9] J. von Neumann and O. Morgenstern, *Theory of games and economic behavior*, 2nd ed. Princeton University Press, 1947.

[10] D. Bernoulli, "Specimen theoriae novae de mensura sortis," *Commentarii Academiae Scientiarum Imperialis Petropolitanae*, vol. 5, 1738, translated by L. Sommer, *Econometrica*, 22: 23–36, 1954.

[11] D. Bell, "One-switch utility functions and a measure of risk," *Management Science*, vol. 34, no. 12, pp. 1416–1424, 1988.

[12] P. Haddawy and S. Hanks, "Utility models for goal-directed decision-theoretic planners," *Computational Intelligence*, vol. 14, no. 3, pp. 392–429, 1998.

[13] R. Bellman, *Dynamic Programming*. Princeton University Press, 1957.

[14] R. Howard, *Dynamic Programming and Markov Processes*. MIT Press, 1960.

[15] R. Howard and J. Matheson, "Risk-sensitive Markov decision processes," *Management Science*, vol. 18, no. 7, pp. 356–369, 1972.

[16] S. Marcus, E. Fernàndez-Gaucherand, D. Hernàndez-Hernàndez, S. Colaruppi, and P. Fard, "Risk-sensitive Markov decision processes," in *Systems and Control in the Twenty-First Century*, C. B. et. al., Ed. Birkhauser, 1997, pp. 263–279.

[17] J. Boyan and M. Littman, "Exact solutions to time-dependent MDPs," in *Advances in Neural Information Processing Systems*, vol. 13, 2000, pp. 1026–1032.

[18] P. Poupart, C. Boutilier, D. Schuurmans, and R. Patrascu, "Piecewise linear value function approximation for factored MDPs," in *Proceedings of the AAAI Conference on Artificial Intelligence*, 2002, pp. 292–299.

[19] Z. Feng, R. Dearden, N. Meuleau, and R. Washington, "Dynamic programming for structured continuous Markov decision problems," in *Proceedings of the International Conference on Uncertainty in Artificial Intelligence*, 2004, pp. 154–161.

[20] L. Li and M. Littman, "Lazy approximation for solving continuous finite-horizon MDPs," in *Proceedings of the AAAI Conference on Artificial Intelligence*, 2005, pp. 1175–1180.

[21] J. Marecki and P. Varakantham, "Risk-sensitive planning in partially observable environments," in *Proceedings of the International Conference on Autonomous Agents and Multiagent Systems*, 2010, pp. 1357–1364.

[22] Y. Liu and S. Koenig, "Risk-sensitive planning with one-switch utility functions: Value iteration," in *Proceedings of the AAAI Conference on Artificial Intelligence*, 2005, pp. 993–999.

[23] ——, "An exact algorithm for solving MDPs under risk-sensitve planning objectives with one-switch utility functions," in *Proceedings of the International Conference on Autonomous Agents and Multiagent Systems*, 2008, pp. 453–460.

[24] ——, "Functional value iteration for decision-theoretic planning with general utility functions," in *Proceedings of the AAAI Conference on Artificial Intelligence*, 2006, pp. 1186–1193.

[25] J. Marecki, S. Koenig, and M. Tambe, "A fast analytical algorithm for solving Markov decision processes with real-valued resources," in *Proceedings of the International Joint Conference on Artificial Intelligence*, 2007, pp. 2536–2541.

[26] J. Marecki and M. Tambe, "Towards faster planning with continuous resources in stochastic domains," in *Proceedings of the AAAI Conference on Artificial Intelligence*, 2008, pp. 1049–1055.

[27] ——, "Planning with continuous resources for agent teams," in *Proceedings of the International Conference on Autonomous Agents and Multiagent Systems*, 2009, pp. 1089–1096.

[28] Mausam and A. Kolobov, *Planning with Markov Decision processes: An AI Perspective*. Morgan and Claypool Publishers, 2012.

[29] C. Boutilier, T. Dean, and S. Hanks, "Decision-theoretic planning: Structural assumptions and computational leverage," *Journal of Artificial Intelligence Research*, vol. 11, pp. 1–94, 1999.

[30] J. Hoey, R. Aubin, and C. Boutilier, "SPUDD: stochastic planning using decision diagrams," in *Proceedings of the International Conference on Uncertainty in Artificial Intelligence*, 1999, pp. 279–288.

[31] E. Hansen and S. Zilberstein, "LAO*: a heuristic search algorithm that finds solutions with loops," *Artificial Intelligence*, vol. 129, pp. 35–62, 2001.

[32] Mausam, E. Benazara, R. Brafman, N. Meuleau, and E. Hansen, "Planning with continuous resources in stochastic domains," in *Proceedings of the International Joint Conference on Artificial Intelligence*, 2005, pp. 1244–1251.

[33] Y. Liu, "Decision-theoretic planning under risk-sensitive planning objectives," Ph.D. dissertation, College of Computing, Georgia Institute of Technology, Atlanta (Georgia), 2005.

[34] W. Powell, *Approximate Dynamic Programming*, 2nd ed. John Wiley and Sons, 2011.

[35] M. Dias, R. Zlot, N. Kalra, and A. Stentz, "Market-based multirobot coordination: A survey and analysis," *Proceedings of the IEEE*, vol. 94, no. 7, pp. 1257–1270, 2006.

[36] P. McAfee and J. McMillan, "Auctions and bidding," *Journal of Economic Literature*, vol. 15, pp. 699–738, 1987.

[37] R. Smith, "The contract net protocol: High level communication and control in a distributed problem solver," *IEEE Transactions on Computers*, vol. C-29, pp. 1104–1113, 1980.

[38] M. Dias, R. Zlot, N. Kalra, and A. Stentz, "Market-based multirobot coordination: A survey and analysis," *Proceedings of the IEEE*, vol. 94, no. 7, pp. 1257–1270, 2006.

[39] M. Lagoudakis, V. Markakis, D. Kempe, P. Keskinocak, S. Koenig, A. Kleywegt, C. Tovey, A. Meyerson, and S. Jain, "Auction-based multi-robot routing," in *Proceedings of the International Conference on Robotics: Science and Systems*, 2005, pp. 343–350.

[40] E. Lawler, J. Lenstra, A. Kan, and D. Shmoys, Eds., *The Traveling Salesman Problem*. John Wiley, 1985.

[41] T. Sandholm, "An implementation of the contract net protocol based on marginal cost calculations," in *Proceedings of the International Workshop on Distributed Artificial Intelligence*, 1993, pp. 295–308.

[42] C. Tovey, M. Lagoudakis, S. Jain, and S. Koenig., "The generation of bidding rules for auction-based robot coordination," in *Multi-Robot Systems: From Swarms to Intelligent Automata*, L. Parker, F. Schneider, and A. Schultz, Eds. Springer, 2005, pp. 3–14.

[43] M. Lagoudakis, P. Keskinocak, A. Kleywegt, and S. Koenig, "Auctions with performance guarantees for multi-robot task allocation," in *Proceedings of the International Conference on Intelligent Robots and Systems*, 2004, pp. 1957–1962.

[44] S. Koenig, C. Tovey, X. Zheng, and I. Sungur, "Sequential bundle-bid single-sale auction algorithms for decentralized control," in *Proceedings of the International Joint Conference on Artificial Intelligence*, 2007, pp. 1359–1365.

[45] S. Koenig, P. Keskinocak, and C. Tovey, "Progress on agent coordination with cooperative auctions [senior member paper]," in *Proceedings of the AAAI Conference on Artificial Intelligence*, 2010.

[46] R. Korf, "Real-time heuristic search," *Artificial Intelligence*, vol. 42, no. 2–3, pp. 189–211, 1990.

[47] I. Nourbakhsh and M. Genesereth, "Assumptive planning and execution: A simple, working robot architecture," *Autonomous Robots*, vol. 3, pp. 49–67, 1996.

[48] S. Koenig, Y. Smirnov, and C. Tovey, "Performance bounds for planning in unknown terrain," *Artificial Intelligence Journal*, vol. 147, no. 1–2, pp. 253–279, 2003.

[49] P. Hart, N. Nilsson, and B. Raphael, "A formal basis for the heuristic determination of minimum cost paths," *IEEE Transactions on Systems Science and Cybernetics*, vol. SCC-4, no. 2, pp. 100–107, 1996.

[50] S. Koenig, M. Likhachev, Y. Liu, and D. Furcy, "Incremental heuristic search in artificial intelligence," *Artificial Intelligence Magazine*, vol. 25, no. 2, pp. 99–112, 2004.

[51] A. Stentz, "The focussed D* algorithm for real-time replanning," in *Proceedings of the International Joint Conference on Artificial Intelligence*, 1995, pp. 1652–1659.

[52] S. Koenig and M. Likhachev, "Fast replanning for navigation in unknown terrain," *Transactions on Robotics*, vol. 21, no. 3, pp. 354–363, 2005.

[53] G. Ramalingam and T. Reps, "An incremental algorithm for a generalization of the shortest-path problem," *Journal of Algorithms*, vol. 21, pp. 267–305, 1996.

[54] A. Barto, S. Bradtke, and S. Singh, "Learning to act using real-time dynamic programming," *Artificial Intelligence*, vol. 72, no. 1–2, pp. 81–138, 1995.

[55] J. Svennebring and S. Koenig, "Building terrain-covering ant robots," *Autonomous Robots*, vol. 16, no. 3, pp. 313–332, 2004.

[56] I. Wagner and A. Bruckstein, "From ants to a(ge)nts: A special issue on ant-robotics," *Annals of Mathematics and Artificial Intelligence*, vol. 31, no. 1–4, pp. 1–5, 2001.

[57] Y. Gabriely and E. Rimon, "Spanning-tree based coverage of continuous areas by a mobile robot," *Annals of Mathematics and Artificial Intelligence*, vol. 31, pp. 77–98, 2001.

[58] ——, "Spanning-tree based coverage of continuous areas by a mobile robot," in *Proceedings of the International Conference on Robotics and Automation*, 2001, pp. 1927–1933.

[59] N. Hazon and G. Kaminka, "Redundancy, efficiency, and robustness in multi-robot coverage," in *Proceedings of the International Conference on Robotics and Automation*, 2005, pp. 735–741.

[60] ——, "On redundancy, efficiency, and robustness in coverage for multiple robots," *Robotics and Autonomous Systems*, vol. 56, no. 12, pp. 1102–1114, 2008.

[61] G. Even, N. Garg, J. Könnemann, R. Ravi, and A. Sinha, "Min-max tree covers of graphs," *Operations Research Letters*, vol. 32, pp. 309–315, 2004.

[62] X. Zheng, S. Koenig, D. Kempe, and S. Jain, "Multi-robot forest coverage for weighted and unweighted terrain," *IEEE Transactions on Robotics*, vol. 26, no. 6, pp. 1018–1031, 2010.