RESEARCH INTERESTS

Ashish Sabharwal
Cornell University

The goal of my research is to *develop scalable and robust reasoning and inference technology* that will allow computers to act intelligently in increasingly complex real-world settings and in competitive and uncertain environments. I am particularly interested Artificial Intelligence techniques centered around Combinatorial Reasoning, Constraints, Probabilistic Inference, Multi-Agent and Adversarial Reasoning, and Optimization, both for traditional computing architectures and for Multi-Core and Cloud Computing platforms. While a significant amount of my work in the past has been driven by traditional applications of these techniques to problems in areas such as planning, scheduling, verification, and diagnosis, the central theme of my current research agenda is the application of AI techniques to the newly emerging, highly inter-disciplinary field of Computational Sustainability. I am actively involved in establishing this new area and in finding innovative ways to inject computational thinking into the process of addressing key environmental and natural resource management challenges of our time.

Inference and optimization engines lie at the core of artificial intelligence, and are often embedded in systems ranging from car diagnosis tools to production planning software in the industry to complex control units in space missions. The vision for this field, at least on the technological side, is well exemplified by the DARPA Grand Challenge, where fully autonomous ground vehicles have relied on automated inference and control methods to successfully navigate over 100 miles of terrain without any human assistance. But *the role of AI methods is not limited to advances in technology and automation alone*—they have a tremendous potential for helping us find better, more effective, and more meaningful ways of addressing broader challenges involving our environment, economy, and society, as I will discuss later in this statement.

Combinatorial reasoning technology, such as for propositional satisfiability (SAT), has witnessed an unprecedented growth. Starting with only a few hundred variable problems in the early 1990's, publicly available SAT solvers can now easily solve many large industrial problems with over 1 million variables and 5 million constraints. This field promises to grow further in richness and impact as we begin to venture beyond basic combinatorial search, to problems such as *inferring properties of complex combinatorial spaces, and reasoning in multi-agent competitive environments.* This requires a shift in technology from NP-complete search problems to significantly more difficult #P- and PSPACE-complete problems. Efficient techniques for these problems are crucial for pushing automated inference engines to a level where they can address complex real-world systems. This poses many fundamental research challenges:

– *Scalability*: How can we overcome the high worst-case computational cost of solving these problems and scale them to the range of millions of variables demanded by real-world applications?

– *Robustness*: How can we obtain performance guarantees and have solutions that are robust against rare but catastrophic failures, such as power grids failures and stock market crashes?

– *Synergy between exact and probabilistic inference*: How can we directly or indirectly use techniques developed for exact inference to solve problems in probabilistic inference, and vice versa?

– *Multi-agent reasoning*: How can we perform strategic decision making that will provably work in the presence of multiple agents and adversaries with competing interests?

– *Representation and balance*: How rich should the problem representation language be? How should we balance brute-force exploration with expensive inference at each step of the search?

– *New computing hardware*: How can complex reasoning engines efficiently and effectively exploit the tremendous computational potential of Cloud Computing and Multi-Core platforms?

I believe addressing these challenges will require a constant flow of ideas between foundational research and implementation, and successful solutions will pave the road to a whole new range of applications. My research approach thus follows *a unique blend of rigorous analysis and principled experimentation*, placing emphasis on studying complex computation processes themselves as natural phenomena, so as to achieve a better understanding of our computational methods and push them beyond their current limits.

In what follows, I have provided a synopsis of my ongoing and past work[1] aligning with four themes. Part A summarizes the exciting, new Computational Sustainability component of my research agenda, while parts B, C, and D are continuations of my earlier work on foundational methods for AI reasoning and inference.

A. Reasoning and Optimization for Computational Sustainability

During the past year and a half, a central focus of my research has been the application of AI techniques, specifically constraint reasoning and optimization, to solve a variety of challenging problems arising in *computational sustainability*—a new, highly inter-disciplinary field that aims to use advanced computational methods as a tool for intelligent decision making concerning the management and allocation of the limited resources on our planet. With the increasingly growing understanding of and concern about the dramatic impact of human development—and ever greater consumption of materials and natural resources—on the earth's delicate balance, I believe it is imperative that computer scientists play a key role in addressing the grand sustainability challenge facing us today.

As a researcher at Cornell's Institute for Computational Sustainability (ICS) founded in 2008, I have been actively involved in creating a vision for the institute, beginning with a successful $10M NSF Expeditions in Computing proposal itself. Besides basic research, I have contributed to nurturing the growth of this new field by fostering interdisciplinary interactions and by co-organizing conferences and workshops to bring together researchers with diverse backgrounds (CompSust09, the 1st International Conference on Computational Sustainability; CROCS-09, the 1st International Workshop on Constraint Reasoning and Optimization for Computational Sustainability).

My research in this area has resulted in a close interaction with environmental agencies such as The Conservation Fund and the National Fish and Wildlife Service, as well as with researchers in other fields such as Resource Economics, Civil and Environmental Engineering, and Materials Science and Engineering. These interactions have been a very enlightening experience, as they have slowly made it apparent how pervasive computer science problems—ranging from those that can be handled through constraint reasoning, search, and optimization to those that ask for novel algorithm design, data mining, human-computer interaction, and machine learning—are in sustainability applications. Unfortunately, the use of computers by researchers and practitioners in sustainability related fields tends to be relatively simplistic, which opens up a tremendous potential for computer scientists to make a significant impact by *injecting computational thinking into sustainability research*. On the other hand, the highly dynamic, complex, and large-scale nature of sustainability problems opens up *new challenges for computer science* itself. Here are some of my current projects exemplifying this:

Wildlife Corridor Design and Bird Conservation: Growing land fragmentation has been found to have a major negative ecological impact on wildlife, and the construction of "corridors" connecting different regions of biological significance has been shown to significantly mitigate this impact. However, the methods currently used to design such corridors are very limited in what they model and capture. Our work has formalized this problem as a budget-constrained utility-optimization problem on undirected graphs with a connectivity constraint [CPAIOR-08/07, in preparation for Conservation Biology J.]. By studying the empirical hardness behavior of this problem under various parameter regimes, we have developed a scalable hybrid, two-phase, MIP-based methodology. Working with researchers in Resource Economics on real-life data obtained for Grizzly Bear population in the U.S. Northern Rockies, with several months of CPU time, we have found optimal or near-optimal corridors at various budget levels connecting Yellowstone, Selmon-Selway, and Northern Continental Divide Ecosystems.

In related work on bird conservation, our group is studying effective ways of conserving the fast diminishing population of Red Cockaded Woodpecker (RCW), a federally endangered species. Working closely with The Conservation Fund, we have formulated this problem as a stochastic constraint optimization problem,

[1] The work discussed here is a result of joint effort with Carlos Ansotegui, Mathieu Blanchette, Paul Beame, Jon Conrad, Carmel Domshlak, Carla Gomes, Willem-Jan van Hoeve, Jörg Hoffmann, Russell Impagliazzo, Henry Kautz, Andrew McCallum, Gilles Pesant, Toniann Pitassi, Ron Raz, Louis-Martin Rousseau, Meinolf Sellmann, and Bart Selman; and (then) students: Josh Buresh-Oppenheim, Matthew Cary, Aron Culotta, Bistra Dilkina, Justin Hart, Ethan Kim, Lukas Kroc, Yuri Malitski, Raghu Ramanujan, Atri Rudra, Jordan Suter, and Erik Vee.

using MIP and pseudo-Boolean frameworks. To address scalability issues, we are employing Cloud Computing methods, using our access to Yahoo!'s "M45" compute cluster. This work provides the first scalable approach for budget-restricted high-quality decision making for land acquisition and cavity implants for RCW and related bird conservation efforts.

Fish Passage Barrier Removal: This work addresses the problem of minimizing the impact of man-made stream barriers on the upstream and downstream movement of migratory fish such as Salmon, which need accessibility to high quality rearing and spawning habitat in side channels or upper tributaries of a river. Various studies have shown that small barriers such as dams, culverts, dikes, levees, floodgates, and weirs have contributed to a significant decrease in fish population (e.g., 40% reduction in Salmon stocks in the Pacific Northwest and over 80% in the Atlantic Northeast). Barrier removal and restoration projects are being pursued in several regions to mitigate this impact. However, the selection of which projects to undertake in the limited budget is often based on simple *scoring and ranking* methods, which are known to be sub-optimal. With interactions with the National Fish and Wildlife Service, we are developing a Mixed Integer Programming based optimization model [preliminary results at CROCS-09] as a scalable technique for finding optimal or near-optimal solutions that make the best possible use of the typically low budget.

Material Discovery: In ongoing work with Material Science researchers, we are using constraint reasoning and machine learning techniques to assist with the discovery of new, rich, and complex materials with certain desirable properties such as "shape memory", with the overall goal of making the best use of our limited natural resources. The central problem here involves analyzing x-ray diffraction patterns in order to identify various structural *phases* in a thin film composition spread experiment. The challenge is to develop effective reasoning and learning methods to overcome the inherent combinatorial explosion and deal with experimental errors and limitations.

Sustainable Community Development and Renewable Energy: The design and operation of zero-impact local communities that can sustain themselves while meeting the economic, social, and environmental aspirations of the residents as well as the broader stakeholders, is finding increasing attention. This yields a complex and dynamic system, with many highly-interacting "agents" with potentially conflicting goals and requirements and several discrete design choices such as viable community size, a balanced portfolio of renewable and non-renewable energy, balanced land use, etc. This ongoing work with Civil, Environmental, and Mechanical Engineers aims to create a scalable mathematical and computational framework for representing and analyzing the costs and impacts of such choices in a realistic manner, through techniques that go beyond traditional Life Cycle Assessment (LCA) analysis and its associated limitations.

B. DEVELOPING SCALABLE AUTOMATED REASONING METHODS

Another focus area of my research is on building fast, practical methods for combinatorial reasoning, often based on non-traditional approaches. The application domains and annual competitions that have been the catalyst for the tremendous growth in SAT solver and related technology have also brought the discipline to the boundary between general scientific research and careful, detailed engineering to excel on specific benchmark problems. A substantial amount of work goes into fine-tuning and exploring variations of techniques already in place. While this undoubtedly has its own merit, it is also crucial to look at problems from a fresh perspective. *Can one address issues regularly encountered by researchers by introducing a fundamentally new way of solving, or even representing, the problem?* This has served as a constant motivating question for my research. Here are four examples of innovative approaches that I have introduced, all of which have pushed the limits of automated reasoning by orders of magnitude:

XOR-Streamlining and Extensions for Model Counting and Sampling: Part of this work was recognized with an Outstanding Paper Award at AAAI-06, the 21^{st} National Conference on Artificial Intelligence. This work addresses the problem of counting and near-uniformly sampling solutions of combinatorial problems [AAAI-06/07, NIPS-06, IJCAI-07, SAT-07] and has applications in areas such as Bayesian inference and probabilistic reasoning. Implemented in a series of state-of-the-art tools (`MBound`, `SampleCount`, `XorSample`), XOR-streamlining is a fundamentally different technique for attacking these #P-hard problems than the traditional ones based on backtracking search and Markov Chain Monte Carlo (MCMC) methods. Inspired by work in complexity theory, this approach introduced the first effective and scalable method for using a complete or local search SAT "solver" essentially *off-the-shelf* to build a "model counter". It also introduced a new framework for obtaining lower and upper bounds with probabilistic correctness guarantees (rather than estimates without guarantees), outperforming previous approaches by several orders of magnitude on challenging problem instances.

Our recent work [CPAIOR-08, ISAIM-08, ANOR J.-08], implemented in tools `BPCount` and `MiniCount`, broadens the scope of these ideas by incorporating message passing techniques from probabilistic inference (namely, Belief Propagation and Survey Propagation) as well as statistical estimation methods. For lower bounds on the model count, this often provides a faster alternative to sampling based techniques, and for upper bounds, this is indeed the first realistically scalable approach with correctness guarantees.

Multi-Core and Cloud Computing Platforms for Inference and Analysis: As multi-core and cloud computing architectures become the dominant high-performance computing platform, a challenge and opportunity for the AI community is to develop effective ways to use these platforms for complex inference and analysis tasks. Our recent work [IJCAI-09, SAT-09] has introduced a low-overhead framework for using multi-core architectures to solve the combinatorial optimization problem of Maximum Satisfiability or MaxSAT, particularly when the instances are "barely" infeasible. The hybrid method uses a loose coupling between a systematic DPLL solver and a local search solver such as Walksat, through a simple shared memory architecture and minimalistic but powerful interaction. For a large suite of challenging industrial problems from the SAT Race-2008 competition, our method is able to find near-optimal solutions in seconds while competing approaches take minutes to hours and still obtain solutions that are far from optimal.

In related ongoing work, we are exploiting the high experimentational capability of cloud computing platforms (specifically, the "M45" compute cluster of Yahoo!) to discover heavy-tailed runtime behavior of local search methods for combinatorial reasoning; such behavior, with its strong implications on solution methods, is now well known for systematic search but has never been clearly observed in local search. We are also using the M45 cluster to develop a scenario-based iterative solution method for a stochastic optimization problem arising in a computational sustainability application discussed earlier.

Dual Formulation for Adversarial Reasoning by QBF solvers: Understanding and avoiding worst-case scenarios is central to mission-critical adversarial reasoning applications where there is no room for error. Quantified Boolean Formula (QBF) reasoning, built upon the success of SAT solvers, is an effective methodology for addressing such applications, but is far less scalable than single-agent SAT reasoning. Implemented as the solver `Duaffle`, our method brought a departure from the commonly used CNF-based representation formalism for QBF reasoning [SAT-07]. By using a novel dual CNF-DNF representation based on a two-player game and planning perspective of QBF domains, this approach brings out the full power of DNF-based "solution learning" techniques and facilitates, for the first time, constraint propagation across quantifiers—a bottleneck for search-based QBF solvers. Our work provided not only an easier and more natural way to formulate adversarial problems in a dual representation but also orders of magnitude improvement in efficiency on many challenging instances.

In ongoing work [partly under review], we have obtained new insights into the nature of adversarial search spaces (specifically, the presence of "early losses" in some domains such as Chess and absence in other domains such as the game of Go) and how differently it influences the effectiveness of two mainstream adversarial reasoning approaches: exhaustive but expensive Minimax search and sampling-based opportunistic UCT-style search.

Symmetry-Breaking and Reasoning Beyond Resolution: The presence of structural symmetry in combinatorial problems artificially increases the effective search space size to explore. Symmetry-breaking approaches are therefore a necessary ingredient for scalable reasoning in such domains. Implemented in a structure-aware SAT solver called `SymChaff`, my technique [AAAI-05, Constraints J.-09] is again a departure from commonly used methods such as "symmetry-breaking constraints" for SAT and CSP. By retaining and exploiting automatically generated contextual information about problem variables, `SymChaff` can achieve as much as (provably) exponential speed-ups over the best alternative approaches. The theme is the same as in the QBF work discussed above: *overcome limitations introduced by traditional* CNF*-based encodings by suitably altering the problem representation formalism.*

C. FORMAL ANALYSIS OF LOGICAL AND PROBABILISTIC INFERENCE ENGINES

Understanding inherent strengths and limitations of various methods used in practice plays a crucial role in further development and successful application of such methods. Following this philosophy, I have created formal frameworks for understanding and analyzing state-of-the-art approaches in satisfiability testing and AI planning, complementing empirical observations about when these approaches work well and when they don't. This has led to new ideas addressing fundamental limitations of known techniques. My work has also explored probabilistic inference methods and the role of problem structure in systematic search. My methodology here has combined a mathematical analysis, often based on proof complexity theory and discrete mathematics, and systematic experimentation. Examples of my work in this direction include:

A Resolution-Based Formal Framework for SAT Solvers: This work [IJCAI-03, SAT-03, JAIR-04], recognized as the Runner-up for the IJCAI-JAIR 5-Year Best Paper Prize for years 2003-2008, introduced the first formal proof complexity based framework for a rigorous analysis of two key techniques often engineered in DPLL-style SAT solver implementations. Specifically, it revealed the inherent power of *clause learning* and *restart* techniques, by relating them to the 'resolution' proof system—the holy grail for all DPLL-based SAT solvers. This laid a much needed foundation for understanding and analyzing SAT solvers, and helped make clause learning and restarts an integral part of the next generation of solvers while also spawning further research and improvement.

Probabilistic Inference and Message Passing: A relatively recent message passing technique originating in the statistical physics community, called Survey Propagation (SP), has turned out to be much more efficient than mainstream DPLL and local search methods for solving very hard random SAT instances. While somewhat related to the more common Belief Propagation (BP) framework, the exact nature of the probabilistic information sought by SP has been elusive, leading to several publications attempting to "explain" SP in rigorous, mathematical terms. Our recent work [in preparation for PNAS, preliminary version at NIPS-08] has developed a new "cluster-centered" approach for explaining SP and deriving SP equations systematically for *any* discrete constraint satisfaction problem or graphical model, the same way one would apply the generic BP framework to such a problem. This cluster-focused approach provides new insights into the solution space structure of combinatorial problems and holds promise for extending the success of SP-style algorithms from random instances to more useful, structured instances.

Our work [CPAIOR-08, ISAIM-08] also demonstrated how such properties of the solution space can be fruitfully exploited to obtain the *number* of solutions of these problems, even in highly structured domains. In related work [UAI-07, SAC-09], we revealed through extensive experimentation the nature of the probabilistic information about the solution space that SP effectively computes. This work was nominated for the Best Student Paper award at UAI-07.

Abstraction in AI Planning: Abstraction is a commonly employed technique, especially in large scale model checking and planning, that attempts to improve efficiency by abstracting away non-critical details of the problem at hand and thus reducing the size of the raw search space. Can abstraction methods really achieve much benefit in AI planning systems? While this empirically does seem to be the case for BDD-based symbolic planners (which essentially perform a fast blind search) and for heuristic planners such as IPP, this work [ICAPS-06, JAIR-09] provided a rather surprising provably *negative* answer for the best-case behavior of Resolution-based optimal planners such as the award winning SATPLAN planner: all of commonly used abstraction methods cannot improve the best case behavior of SATPLAN under several encodings typically used in practice. At a high level, this showed that the "informedness" of the search method must compete with the informedness of the abstraction heuristic, providing new insights into the design of abstraction techniques. This work was nominated for the Best Paper Award at the ICAPS-06 conference.

Hardness Profiles and Problem Structure: Constraint satisfaction problems often exhibit an intriguing pattern: an abrupt phase transition from being feasible to being infeasible as a key problem parameter is varied. This work [CPAIOR-07/08], with direct application to "wildlife corridor" design for grizzly bears in the U.S. Northern Rockies discussed earlier, empirically revealed for the first time such phenomena—and a corresponding "easy-hard-easy" pattern—for problems that combine both constraint satisfaction and optimization aspects. In ongoing work, we have discovered—for the first time—heavy tail runtime distribution patterns in *local search* solvers.

In a related direction [CP-07, ISAIM-08], our work has brought to light the fundamental strength of the notion of propagation-based "backdoor sets" used to characterize real-world structure in combinatorial problems and explain the astonishing scalability of SAT solvers on structured industrial benchmarks. Specifically, we showed that *dynamic*, propagation based backdoors used by these solvers can be exponentially smaller, and thus more powerful, than those based on *static*, syntactic classes such as 2-CNF or Horn SAT used in certain formal analyses, and also provably harder to compute in the worst case assuming NP $\neq$ coNP.

Building upon this work further [SAT-09], we extended the notion of backdoors to capture *learning during search*—a key ingredient of many successful state-of-the-art constraint solvers—and showed that learning again can result in exponentially more compact backdoors. Finally, we extended the notion of backdoors, defined originally for constraint satisfaction problems, to general *optimization* problems and demonstrated that interesting MIP (mixed integer programming) optimization problems do have surprisingly short backdoor sets [CPAIOR-09]. In fact, contrary to the general intuition about computational hardness in optimization, we found that for some problems, backdoors for proving optimality were significantly smaller than those for finding an optimal solution.

D. FOUNDATIONAL ISSUES: ALGORITHM DESIGN AND PROOF COMPLEXITY

The final theme of my research involves addressing foundational issues underlying automated reasoning systems. Specifically, I design efficient algorithms for constraint solvers and characterize the strength of various "proof systems". This has resulted in the first polynomial time algorithms in some cases, and NP-completeness or hardness of approximation results in other. Two examples of my work in this area are:

Filtering Algorithms for Special Constraints: This work [CP-06, Constraints J.-09] was recognized with the Best Paper Award at CP-06, the 12^{th} International Conference on Principles and Practice of Constraint Programming. It introduced the first polynomial time filtering algorithm for a combinatorial constraint (the "sequence" constraint) that appears frequently in scheduling and design automation problems, such as in a car manufacturing pipeline. This resolved a question that had been open for 10 years in the Constraint Programming (CP) community.

In related work [ModRef-07, CPAIOR-08, ongoing], our algorithms have revealed the exponential memory and runtime savings that higher-level *set-based representations* of constraints can achieve. This work provided some of the first polynomial time filtering algorithms for constraints in this higher-level representation.

Resolution Complexity and Hardness of Approximation: This work in the field of proof complexity theory [Complexity-01, Computational Complexity J.-07] studied the Resolution proof system in semi-structured domains. It showed that almost all instances of some interesting co-NP complete graph problems require exponential size Resolution proofs of infeasibility, even to approximate within significant factors, thus providing a large family of structured formulas that are exponentially hard for the Resolution proof system. The methodology involved a blend of combinatorial and probabilistic analysis, and expansion properties of random graphs. The work also showed that a natural class of approximate optimization algorithms for these problems must fail to provide good approximations on almost all problem instances. In related work [FOCS-02, SIAM J. Computing-04], we proved that even stronger proof systems, such as bounded-depth Frege systems described in many logic texts, require exponential size proofs even for very weak pigeonhole formulas, strengthening previously known results in this area.