SHEILA A. MCILRAITH

Department of Computer Science, University of Toronto
Canada CIFAR AI Chair (Vector Institute)
Associate Director, Schwartz Reisman Institute for Technology and Society
Email: sheila@cs.toronto.edu; Home: https://www.cs.toronto.edu/∼sheila
Google Scholar

EDUCATION

Ph.D. (Computer Science), University of Toronto, Spring, 1997.
Thesis title: Towards a formal account of diagnostic problem solving.
Supervisor: Prof. Raymond Reiter.

EMPLOYMENT HISTORY

March 2020 – present	Associate Director (July 2021) and Research Lead Schwartz Reisman Institute for Technology and Society University of Toronto
October 2019 – 2021	Visiting Consulting Professor Samsung Research Laboratory Canada, Toronto, Canada
September 2019 – present	Canada CIFAR AI Chair and Faculty Member Vector Institute for Artificial Intelligence, Toronto, Canada
July 2012 – present	Professor Department of Computer Science, University of Toronto
July 2011 – June 2013	Vice Chair Department of Computer Science, University of Toronto
Nov. 2003 – June 2012	Associate Professor Department of Computer Science, University of Toronto Tenure awarded: July, 2007
Mar. 2004 – Feb. 2005	Visiting Associate Professor Department of Computer Science, Stanford University
Jan. 1998 – Sept. 2003	(Senior as of 2002) Research Scientist Knowledge Systems Laboratory Department of Computer Science, Stanford University
Feb. 1997 – Dec. 1997	Postdoctoral Fellow Department of Computer Science, Stanford University
Feb. 1997 – Dec. 1997	Visiting Scholar Systems and Practices Laboratory Xerox Palo Alto Research Center (PARC), Palo Alto, CA, USA

HONOURS

2025	Schmidt Sciences AI2050 Senior Fellow. (7 senior fellows appointed internationally in 2025.)
2025	D2L Innovation Award in Teaching and Learning. National award recognizing excellence in collaborative post-secondary teaching and learning for the University of Toronto Embeded Ethics Education Initiative (E3I) (team award with D. Horton, D. Liu, S. Coyne)
2024	The Canadian AI Association (CAIAC) Lifetime Achievement Award. Recognizing outstanding research excellence in AI.
2024	Northrop Frye Award, a University of Toronto Award of Excellence for the University of Toronto Embedded Ethics Education Initiative (E3I) (team award with D. Horton, D. Liu, S. Coyne)
2023	IJCAI-JAIR Best Paper Prize, one outstanding paper published in the Journal of AI Research (JAIR) in last 5 years, based on its significance and presentation quality. (with R. Toro Icarte, T. Klassen, R. Valenzano)
2022	ICAPS Influential Paper Award, test of time award honoring significant and influential paper from the International Conference on Automated Planning and Scheduling, 10 years prior. (with C. Muise and C. Beck)
2022	AI Risk Analysis Award. NeurIPS 2022 ML Safety Workshop. (multiple awarded) (with P. Alzadeh Alamdari and T. Klassen)
2020	Best Paper Award. Knowledge Representation & Reasoning Meets Machine Learning Workshop @ NeurIPS2020 (with L. Illanes, X. Yan, R. Toro Icarte)
2020	Named to "World's 50 Most Renowned Women in Robotics" by Analytics Insight.
2019	Fellow of the Association for Computing Machinery (ACM).
2019	CIFAR Canada AI Chair.
2018	Discovery Accelerator Supplement Grant. Natural Sciences and Engineering Research Council of Canada (NSERC) (125 awarded nationally in all science and engineering disciplines.)
2018	Best Demo Award. International Conference on Automated Planning and Scheduling (ICAPS). (with A. Camacho, J. Baier, and C. Muise)
2015	President, KR Inc. (2015-2016) International not for profit scientific foundation fostering research and communication in knowledge representation and reasoning.
2011	SWSA 10-year Award, recognizing the highest impact paper from the International Semantic Web Conference, 10 years prior. (with A. Ankolekar, M. Burstein, J. Hobbs, O. Lassila, D. Martin, S. Narayanan, M. Paolucci, T. Payne, K. Sycara, and H. Zeng)

2011	Fellow Association for the Advancement of Artificial Intelligence (AAAI). (6 fellows appointed internationally in 2011.)
2011	Best Papers of RE'10 (2010 International Requirements Engineering Conference) (with S. Liaskos, S. Sohrabi, and J. Mylopoulos) (5 recognized.)
2010	Top cited paper in 25 years of IEEE Intelligent Systems[1]. (with T.C. Son, and H. Zeng)
2010	University of Toronto "Women Making a Difference" Distinction.
2007	ACM SIGSOFT Distinguished Paper Award. 22nd IEEE/ACM International Conference on Automated Software Engineering (with Y. Wang, Y. Yu, and J. Mylopoulos) (3 awarded. 12% acceptance rate to conference.)
2007	Discovery Accelerator Supplement Grant. Natural Sciences and Engineering Research Council of Canada (NSERC) (50 awarded nationally in all science and engineering disciplines.)
2006	Early Researcher Award. Ontario Ministry of Research and Innovation (104 awarded in all disciplines in Ontario.)
2006	Distinguished Performance in Satisficing Planning (Qualitative Preferences Track). Fifth International Planning Competition (with J. Baier and F. Bacchus)
2004	NSERC University Faculty Award (UFA). (26 awarded nationally in all science and engineering disciplines.)
1997	NSERC Postdoctoral Fellowship.
1996	Best Student Paper Award. Seventh International Workshop on Principles of Diagnosis

[1] Feng et al., *A Bibliographic Analysis of IEEE Intelligent Systems Publications*. IEEE Intelligent Systems 25(6): 59-66 (2010).

Individual honours bestowed upon my students (excluding scholarships, and awards noted above)

2024	Maayan Shvo (PhD student) 2024 Best Doctoral Dissertation Award, Honorable Mention. International Conference on Automated Planning and Scheduling (ICAPS)
2023	Alberto Camacho (PhD student) 2023 Best Doctoral Dissertation Award. International Conference on Automated Planning and Scheduling (ICAPS)
2023	Rodrigo Toro Icarte (PhD student) 2023 Best Doctoral Dissertation Award, Honorable Mention. International Conference on Automated Planning and Scheduling (ICAPS)
2015	Christian Muise (PhD student) 2015 Best Doctoral Dissertation Award. International Conference on Automated Planning and Scheduling (ICAPS)
2011	Jorge Baier (PhD student) 2011 Best Doctoral Dissertation Award, Honorable Mention International Conference on Automated Planning and Scheduling (ICAPS)
2010	Shirin Sohrabi (PhD student) Best Doctoral Consortium Paper Award International Semantic Web Conference (ISWC)
2010	Christian Fritz (PhD student) 2010 Best Doctoral Dissertation Award, Honorable Mention International Conference on Automated Planning and Scheduling (ICAPS)
2007	Christian Fritz (PhD student) Best Doctoral Consortium Paper Award International Conference on Automated Planning and Scheduling (ICAPS)

PROFESSIONAL ACTIVITIES

Professional Affiliations

Fellow and Member of the Association for Computing Machinery (ACM).
Fellow and Member Association for the Advancement of Artificial Intelligence (AAAI).
Member Canadian Artificial Intelligence Association (CAIAC).
Faculty Member, Vector Institute for Artificial Intelligence, Toronto, Canada.
Affiliate Member, Robotics Institute, University of Toronto, Toronto, Canada.

Distinguished Administrative Service

Canadian AI/CRV 2026 Advisory Board (09/2025-present).
Canadian AI Safety Institute (CAISI) Research Council Member (2025-present)
Associate Director, Schwartz Reisman Institute for Technology and Society (2021 – present).
Research Lead, Schwartz Reisman Institute for Technology and Society (2020 – present).
Board of Directors, KR Inc. (2021 – present).
Chair, Standing Committee One Hundred Year Study on Artificial Intelligence (AI100) (2025 – present).
Chair-Elect, Standing Committee One Hundred Year Study on Artificial Intelligence (AI100) (2023 – 2025).
Standing Committee, One Hundred Year Study on Artificial Intelligence (AI100) (2018 – present).
Chair ACM/AAAI Allen Newell Award Committee (2021).
ACM/AAAI Allen Newell Award Committee (2018 - 2021).
AAAI Fellows Selection Committee (2019 - 2021).
KR Inc. Advisory Board (2018 – present).
Past-President, KR Inc. (2016 – 2018).
President, KR Inc. (2015 – 2016).
AAAI Conference Committee (2018 – 2021).
Advisory Committee, International Joint Conference on Artificial Intelligence, IJCAI (2016).
AAAI Executive Council Member (2007 – 2010).
AAAI Presidential Panel on Long-term AI Futures (2008 – 2009).

Editorial Service

Associate Editor, Journal of AI Research (JAIR). (2015 – 2025)
Editorial Board, AI Magazine, (2014 – 2018)
Associate Editor, Artificial Intelligence Journal (AIJ). (2004 – 2012)
Guest Editor, Artificial Intelligence Journal, Special Issue in Honor of John McCarthy. (2011)
Guest Editor, Journal of Logic and Computation, Special Issue on Logical Formalization of Commonsense Reasoning. (2007)
Editorial Board of Journal of Artificial Intelligence Research (JAIR). (2003–2005)

Selected Conference (Co-)Chairs

ICAPS-24: International Conference on Automated Planning and Scheduling (ICAPS), June, 2024.

Selected Programme (Co-)Chairs

AAAI-18: Thirty-Second AAAI Conference on Artificial Intelligence, February, 2018.
KR2012: Thirteenth International Conference on Principles of Knowledge Representation and Reasoning, June, 2012.
CommonSense 2005: Seventh International Symposium on Logical Formalizations of Commonsense Reasoning, Corfu, Greece, May, 2005.
ISWC2004: Third International Semantic Web Conference, Hiroshima, Japan, November, 2004.
ESSW2002: E-Services and the Semantic Web Workshop, Toronto, Canada, May, 2002.
CogRob2002: Third International Workshop on Cognitive Robotics, Edmonton, Canada, July, 2002.
DX'2001: Twelfth International Workshop on Principles of Diagnosis, Sansicario, Italy, March 2001.
Hybrid Systems and AI: Modeling, Analysis and Control of Discrete + Continuous Systems, 1999 AAAI Spring Symposium, Stanford, USA, March 1999.

Selected Conference Organizing Committees

IJCAI 2025: Advisory Board

AAAI17: Co-chair AAAI Women's Mentoring Breakfast

KR2004: Workshops Chair, Ninth International Conference on the Principles of Knowledge Representation and Reasoning.

KR2004: Doctoral Consortium Chair, Ninth International Conference on the Principles of Knowledge Representation and Reasoning.

ISWC2003: Workshops Chair, Second International Semantic Web Conference.

Selected Workshop Organizing Committees

Co-chair, Dagstuhl Seminar on Cognitive Robotics, 2022.

Co-chair, Dagstuhl Seminar on Epistemic Planning, 2017.

ICAPS-2017 Workshop on Generalized Planning

AAAI-2015 Workshop AI for Cities

AAAI-2014 Workshop on Semantic Cities

AAAI-2011 Workshop on Generalized Planning

ICAPS-2009 Workshop on Generalized Planning: Macros, Loops, Domain Control Knowledge

AAAI-2006 Spring Symposium on Formalizing and Compiling Background Knowledge and its Application to KR and QA

AAAI-2005 Workshop on Exploiting Planning and Scheduling for Web Services, Grid and Autonomic Computing

WWW-2005 Workshop on Web Service Semantics: Towards Dynamic Business Integration; ICAPS-2004 Workshop on Planning and Scheduling for Web and Grid Services

Selected (Senior) Area Chair/Senior Programme Committees

AAAI: Conference on Artificial Intelligence (2006, 2007, 2010, 2011, 2015, 2016, 2017–2026, inclusive).

ICAPS: International Conference on AI Planning and Scheduling (2011, 2013, 2014, 2017, +).

IJCAI: International Joint Conference on Artificial Inteligence (2009, 2011, 2013, 2015, 2016, 2017, 2019, 2020).

ISWC: International Semantic Web Conference (2005, 2009).

KR: International Conference on Principles of Knowledge Representation and Reasoning (2014, 2018, 2020, 2021, 2022).

Selected Programme Committees

AAMAS: Autonomous Agents and Multi-Agent Systems (2023 Blue Sky)

AI: Canadian AI Conference (2013)

AAAI: American National Conf. on Artificial Intelligence (1994, 1998, 1999, 2000, 2002, 2005, 2008).

HSCC: Hybrid Systems: Computation and Control (2002).

ICAPS: International Conference on Automated Planning and Scheduling (2007, 2008, 2009, 2010, 2015, 2019).

IJCAI: International Joint Conference on Artificial Intelligence (2005, 2007).

ISWC: International Semantic Web Conference (2002).

KR: International Conference on Principles of Knowledge Representation and Reasoning (2002, 2006, 2008, 2010).

WWW: International World Wide Web Conference (2002, 2004, 2005, 2006).

CogRob2004: Fourth International Cognitive Robotics Workshop (2004, 2014).

Common Sense: International Symposium on Logical Formalizations of Commonsense Reasoning (2001, 2009).

DX: International Workshop on Principles of Diagnosis (1993, 1994, 1997, 1998, 2001, 2014).

DX/QR: Joint International Workshop on Principles of Diagnosis and Qualitative Reasoning (1999).

SaTML: IEEE Conference on Secure and Trustworthy Machine Learning (2022)

SWWS'01: Semantic Web Workshop (2001).

NMR: International Workshop on Nonmonotonic Reasoning (2000)

Journal Reviewing

Artificial Intelligence;
Annals of Mathematics and Artificial Intelligence;
Computational Intelligence;
Journal of Artificial Intelligence Research;
Journal of Distributed and Parallel Databases;
Journal of Logic Programming;
Journal of Web Semantics.

Advisory Boards and Steering Committees

Stanford One Hundred Year Study on Artificial Intelligence (AI100). Standing Committee. (2018 – present)
Principles of Knowledge Representation and Reasoning Inc. (KR Inc.), Steering Committee. (2004 – 2006; 2010 – 2018)
Syngli (formerly Neurecall). Advisor. (2015 – 2019)
Commonsense Reasoning Org. Advisory Board. (2008 – present)
Semantic Web Service Initiative (SWSI). Advisory Board. (2002 – 2007)

Standardization Efforts

OWL Services Coalition (formerly DAML Services Coalition).
Founder and Member.
2001-2008
A coalition of North American researchers from academia and industry developing a standardized Web service ontology for the World Wide Web in the Semantic Web language OWL.

The result, OWL-S, is an acknowledged World Wide Web Consortium (W3C) member submission (November 2004).

`http://www.w3.org/Submission/OWL-S/`

Semantic Web Service Language (SWSL) Committee.
Founder and Member.
2002 – 2008
SWSL is a coalition of international researchers from academia and industry developing a standardized first-order logic Web service ontology and an associated execution environment for the World Wide Web.

The result, the Semantic Web Services Framework (SWSF) is an acknowledged World Wide Web Consortium (W3C) member submission (September 2005).

`http://www.w3.org/Submission/SWSF/`

REFEREED PUBLICATIONS
Google Scholar Citation Count: 23,594, H-index: 65 (Aug, 2025)

REFEREED JOURNAL ARTICLES

Bengio, Y., Hinton, G., Yao, A., Song, D., Abbeel, P., Darrell, T., Harari, Y. N., Zhang, Y., Xue, L., Shalev-Shwartz, S., Hadfield, G., Clune, J., Maharaj, T., Hutter, F., Baydin, A. G., McIlraith, S., Gao, Q., Acharya, A., Krueger, D., Dragan, A., Torr, P., Russell, S., Kahneman, D., Brauner, J., and Mindermann, S. Managing extreme AI risks amid rapid progress. *Science* 384 (6698), 842-845, 2024.

Horvitz, E., Conitzer, V., McIlraith, S., and Stone, P. Now, Later, and Lasting: Ten Priorities for AI Research, Policy, and Practice. *Communications of the ACM* (CACM), 67 (6), 39-40, 2024.

Majedi, M., McClure, E., Wald, B., Horton, D., and McIlraith, S.A. Embedded Ethics: Pandemic Contact Tracing and Ethical Trade-offs. In ACM EngageCSEdu, 2024. ACM (4 pages + extensive online materials)

Majedi, M., McClure, E., Wald, B., Horton, D., and McIlraith, S.A. Embedded Ethics: Pandemic Exposure Notification Systems and Giving Ethical Justifications. In ACM EngageCSEdu, 2024. ACM (4 pages + extensive online materials)

Toro Icarte, R., Klassen, T. Q., Valenzano, R., Castro, M. P., Waldie, E., and McIlraith, S. A. Learning Reward Machines: A Study in Partially Observable Reinforcement Learning. *Artificial Intelligence* (AIJ), 323: 103989. 2023.

Shati, P., Cohen, E., and McIlraith, S. A. SAT-based optimal classification trees for non-binary data. *Constraints*, 28(2): 166-202. 2023.

Muise, C., Belle, V., Felli, P., McIlraith, S.A., Miller, T., Pearce, A., Sonenberg, L. Efficient multi-agent epistemic planning: Teaching planners about nested belief. *Artificial Intelligence* (AIJ). 302: 103605, 2022.

Baier, J., and McIlraith, S. A. Knowledge-based programs as building blocks for planning. *Artificial Intelligence* (AIJ) 303: 103634, 2022.

Toro Icarte, R., Klassen, T., Valenzano, R., McIlraith, S. Reward Machines: Exploiting Reward Function Structure in Reinforcement Learning. *Journal of Artificial Intelligence Research* (JAIR) 73: 173-208, 2022.

Muise, C., Beck, J.C., McIlraith, S. Optimal Partial-Order Plan Relaxation via MaxSAT. *Journal of Artificial Intelligence Research* (JAIR), 57:113–149, 2016.

Razavi, N., Farzan, A., and McIlraith, S. Generating Effective Tests for Concurrent Programs via AI Automated Planning Techniques. *International Journal on Software Tools for Technology Transfer*, 16(1):49-65, 2014.

Liaskos, S., McIlraith, S., Sohrabi, S., and Mylopoulos, J. Representing and Reasoning about Preferences in Requirements Engineering. *Requirements Engineering*, Special Issue: Best Papers of RE'10, 16(3): 227-249, 2011.

Bienvenu, M., Fritz, C., and McIlraith, S. Specifying and Generating Preferred Plans. *Artificial Intelligence* (AIJ), 175(7-8): 1308-1345, 2011.

Morgenstern, L. and McIlraith, S. John McCarthy's Legacy. *Artificial Intelligence* (AIJ), 175(1):1-24, 2011.

Baier, J., Bacchus, F., and McIlraith, S. A Heuristic Search Approach to Planning with Temporally Extended Preferences. *Artificial Intelligence* (AIJ), 173(5-6):593-618, 2009.

Wang, Y., McIlraith, S., Yu, Y., and Mylopoulos, J. Monitoring and Diagnosing Software Reqluirements. *Journal of Automated Software Engineering*, 16(1):3-35, 2009.

Baier, J., and McIlraith, S. Planning with Preferences. *AI Magazine*, 29(4):25-36, 2009.

Gruninger, M., Hull, R., and McIlraith, S. A Short Overview of FLOWS: A First-Order Logic Ontology of Web Services. *IEEE Data Engineering Bulletin*, 31(3):3-7, 2008.

Martin, D., Burstein, M., McDermott, D., McIlraith, S., Paolucci, M., Sycara, K., McGuinness, D., Sirin, E., Srinivasan, N. Bringing Semantics to Web Services with OWL-S. *World Wide Web*, Special Issue: Recent Advances in Web Services, 10(3):243–277, 2007.

Son, T.C., Baral, C., Nam, T.H., McIlraith, S. Domain-Dependent Knowledge in Answer Set Planning. *ACM Transactions on Computational Logic* (TOCL), 7(4), 2006.

Amir, E. and McIlraith, S. Partition-Based Logical Reasoning for First-Order and Propositional Theories. *Artificial Intelligence* (AIJ), 162(1–2):49–88, 2005.

Choueiry, B., Iwasaki, Y., and McIlraith, S. Towards a Practical Theory of Reformulation for Reasoning about Physical Systems. *Artificial Intelligence* (AIJ), 162(1–2):145–204, 2005.

Narayanan, S. and McIlraith, S. Analysis and Simulation of Web Services. *Computer Networks*, 42(2003):675-693, Elsevier Science B.V., 2003.

Bryson, J., Martin, D., McIlraith, S. and Stein, L. Toward Behavioral Intelligence in the Semantic Web. *IEEE Computer, Special Issue on Web Intelligence*, 35(11):48–55, November, 2002.

McIlraith, S., Son, T.C. and Zeng, H. Semantic Web Services. *IEEE Intelligent Systems, Special Issue on the Semantic Web*, 16(2):46–53, March/April, 2001.

McIlraith, S. Integrating Actions and State Constraints: A Closed-Form Solution to the Ramification Problem (Sometimes). *Artificial Intelligence* (AIJ), 116(1-2):87-121, 2000.

McIlraith, S. Further Contributions to Characterizing Diagnosis. *Annals of Mathematics and Artificial Intelligence*, 11(1–4):137–167, 1994.

McIlraith, S. Qualitative Data Modeling: Application of a Mechanism for Interpreting Graphical Data. *Computational Intelligence*, 5(2):111–120, 1989.

REFEREED CONFERENCE PUBLICATIONS

AAAI: *Conference on Artificial Intelligence*
AAMAS: *International Conference on Autonomous Agents and Multiagent Systems*
AI: *Canadian Conference on Artificial Intelligence*
CP: *International Conference on Principles and Practice of Constraint Programming*
COLM: *Conference on Language Modeling*
ICAPS: *International Conference on Automated Planning and Scheduling*
ICLR: *International Conference on Learning Representations*
ICML: *International Conference on Machine Learning*
IJCAI: *International Joint Conference on Artificial Intelligence (previously biennial)*
IROS: *IEEE/RSJ International Conference on Intelligent Robots and Systems*
ISWC: *International Semantic Web Conference*
KR: *International Conferences on Knowledge Representation and Reasoning (previously biennial)*
RLC: *Reinforcement Learning Conference*
RLDM: *The Multi-disciplinary Conference on Reinforcement Learning and Decision Making* [*]
NeurIPS: *Conference on Neural Information Processing*
SAT: *International Conference on Theory and Applications of Satisfiability*
SoCS: *International Symposium on Combinatorial Search*
SIGCSE: *ACM Technical Symposium on Computer Science Education*
UAI: *Conference on Uncertainty in Artificial Intelligence*
WWW: *World Wide Web Conference*
[*] - *high acceptance rate*

Lifshitz, S., McIlraith, S.A., and Du, Y., Multi-Agent Verification: Scaling Test-Time Compute with Multiple Verifiers, *COLM*, To appear, 2025.

Varricchione, G., Klassen, T.Q., Alechina, N. Dastani Logan, B., and McIlraith, S.A. Pushdown Reward Machines for Reinforcement Learning. *KR*, To appear, 2025.

Li, A. C., Chen, Z., Klassen, T. Q., Vaezipoor, P., Toro Icarte, R. , and McIlraith, S.A. Reward Machines for Deep RL in Noisy and Uncertain Environments. *NeurIPS*, pages 110341–110368, 2024.

Alamdari, P. A., Klassen, T. Q., Creager, E., and McIlraith, S. A. Remembering to Be Fair: On Non-Markovian Fairness in Sequential Decision Making. *ICML*, pages 906–920, 2024.

Shati, P., Cohen, E., and McIlraith, S.A. Neural Sequence Generation with Constraints via Beam Search with Cuts: A Case Study on VRP. *SoCS*, pages 118-126, 2024.

Muise, C., McIlraith, S. A., and Beck, J. C. PRP Rebooted: Advancing the State of the Art in FOND Planning. *AAAI*, pages 20212-20221, 2024.

Horton, D., Liu, D., McIlraith, S. A., Coyne, S., and Wang, N. Do Embedded Ethics Modules Have Impact Beyond the Classroom?. *SIGCSE*, pages 533-539, 2024.

Lifshitz, S., Paster, K., Chan, H., Ba, J., and McIlraith, S. STEVE-1: A Generative Model for Text-to-Behavior in Minecraft. *NeurIPS*, 2023. (Spotlight).

Klassen, T. Q., Muise, C., and McIlraith, S. A. Planning with Epistemic Preferences. *KR*, pages 752-756, 2023.

Wang, A., Li, A. C., Klassen, T. Q., Toro Icarte, R., and McIlraith, S. A. Learning Belief Representations for Partially Observable Deep RL. *ICML*, pages 35970-35988, 2023.

Klassen, T. Q., Alamdari, P. A., and McIlraith, S. A. Epistemic Side Effects: An AI Safety Problem. *AAMAS*, pages 1797-1801, 2023.

Shati, P., Cohen, E., and McIlraith, S. A. SAT-Based Learning of Compact Binary Decision Diagrams for Classification. *CP*, pages 1-19, 2023.

Shati, P., Cohen, E., and McIlraith, S. A. Optimal Decision Trees For Interpretable Clustering with Constraints. *IJCAI*, pages 2022-2030, 2023.

Horton, D., Liu, D., McIlraith, S. A., and Wang, N. Is More Better When Embedding Ethics in CS Courses? *SIGCSE*, pages 652-658, 2023.

Paster, K., McIlraith, S. A., and Ba, J. You Can't Count on Luck: Why Decision Transformers and RvS Fail in Stochastic Environments. *NeurIPS*, 2022.

Tuli, M., Li, A. C., Vaezipoor, P., Klassen, T. Q., Sanner, S., and McIlraith, S. A. Learning to Follow Instructions in Text-Based Games. *NeurIPS*, 2022.

Shvo, M., Hari, R., O'Reilly, Z., Abolore, S., Wang, S. N., and McIlraith, S. A. Proactive Robotic Assistance via Theory of Mind. *IROS*, 2022.

Shvo, M., Klassen, T., McIlraith, S., Resolving Misconceptions about the Plans of Agents via Theory of Mind. *ICAPS*, pages 719-729, 2022.

Alizadeh Alamdari, P., Klassen, T. Toro Icarte, R., McIlraith, S. Be Considerate: Avoiding Negative Side Effects in Reinforcement Learning. *AAMAS*, pages 18-26, 2022.

Klassen, T., McIlraith, S., Muise, C., Xu, J. Planning to Avoid Side Effects. *AAAI*, pages 9830-9839, 2022.

Horton, D., McIlraith, S., Wang, N., Majedi, M., McClure, E., Wald, B. Embedding Ethics in Computer Science Courses: Does it Work? *SIGCSE*, pages 481-487, 2022.

Alizadeh Alamdari, P., Klassen, T. Toro Icarte, R., McIlraith, S. Be Considerate: Avoiding Negative Side Effects in Reinforcement Learning (Extended Abstract), *RLDM*, pages 35-39, 2022.

Vaezipoor, P. Li, A., Toro Icarte, R., McIlraith, S. Achieving Zero-Shot Task Generalization with Formal Language Instructions *RLDM*, pages 627-631, 2022.

Shvo, M., Li, A., Toro Icarte, R., McIlraith, S., Interpretable Sequence Classification via Discrete Optimization. *AAAI*, pages 9647-9656, 2021.

Shvo, M., Hu, Z., Toro Icarte, R., Mohomed, I., Jepson, A., McIlraith, S. AppBuddy: Learning to Accomplish Tasks in Mobile Apps via Reinforcement Learning. *Canadian Conference on AI*, 2021

Pouya Shati, Eldan Cohen, Sheila A. McIlraith: SAT-Based Approach for Learning Optimal Decision Trees with Non-Binary Features. CP 2021: 50:1-50:16

Paster, K., McIlraith, S., Ba, J. Planning from Pixels using Inverse Dynamics Models. *ICLR*, 2021.

Vaezipoor, P. Li, A., Toro Icarte, R., McIlraith, S. LTL2Action: Generalizing LTL Instructions for Multi-Task RL. *ICML*, pages 10497-10508, 2021.

Cohen, E., Valenzano, R., McIlraith, S. Type-WA*: Using Exploration in Bounded Suboptimal Planning. *IJCAI*, pages 4047-4053, 2021

Illanes, L., Yan, X., Toro Icarte, R., and McIlraith, S. Symbolic Plans as High-Level Instructions for Reinforcement Learning. *ICAPS*, pages 540-550, 2020.

Shvo, M., and McIlraith, S. Active Goal Recognition, *AAAI*, pages 9957-9966, 2020.

Shvo, M., Klassen, T., Sohrabi, S., and McIlraith, S., Epistemic Plan Recognition. *AAMAS*, pages 1251-1259, 2020.

Klassen, T., McIlraith, S., and Levesque, H., Changing Beliefs about Domain Dynamics in the Situation Calculus. *KR*, pages 572-581, 2020.

Toro Icarte, R., Waldie, E., Klassen, T., Valenzano, R., Castro, M., McIlraith, S., Learning Reward Machines in Partially Observable Environments (Extended Abstract). *KR* (Recently Published Track), 2020.

Toro Icarte, R., Waldie, E., Klassen, T., Valenzano, R., Castro, M., McIlraith, S., Learning Reward Machines in Partially Observable Environments. *NeurIPS* (Spotlight), pages 15497-15508, 2019.

Toro Icarte, R., Illanes, L., Castro, M., Cire, A., McIlraith, S., Beck, C. Training Binarized Neural Networks using MIP and CP. *CP*, pages 401-417, 2019.

Camacho, A., Toro Icarte, R., Klassen, T., Valenzano, R., McIlraith, S. LTL and Beyond: Formal Languages for Reward Function Specification in Reinforcement Learning. *IJCAI*, 6065-6073, 2019.

Camacho, A., McIlraith, S. Strong Fully Observable Non-Deterministic Planning with LTL and LTL-f Goals. *IJCAI*, pages 5523-5531, 2019.

Camacho, A., McIlraith, S. Learning Interpretable Models in Linear Temporal Logic. *ICAPS*, pages 621-630, 2019.

Camacho, A., Bienvenu, M., and McIlraith, S. Towards a Unified View of AI Planning and Reactive Synthesis. *ICAPS*, pages 58-67, 2019.

Toro Icarte, R., Waldie, E., Klassen, T., Valenzano, R., Castro, M., McIlraith, S., Searching for Markovian Subproblems to Address Partially Observable Reinforcement Learning, *RLDM*, pages 22-26, 2019. [*]

Illanes, L., Yan, X., Toro Icarte, R., McIlraith, S. Symbolic Planning and Model-Free Reinforcement Learning: Training Taskable Agents. *RLDM*, pages 191-195, 2019. [*]

Camacho, A., McIlraith, S. Deep Reactive Synthesis of Linear Temporal Logic Specifications. *RLDM*, 2019. (poster only) [*]

Illanes, L., McIlraith, S. Generalized Planning Via Abstraction: Arbitrary Numbers of Objects. *AAAI*, pages 7610-7618, 2019.

Camacho, A., Muise, C., Baier, J., McIlraith, S. LTL Realizability via Safety and Reachability Games. *IJCAI*, pages 4683-4691, 2018.

Toro Icarte, R., Klassen, T., Valenzano, R., McIlraith, S. Using Reward Machines for High-Level Task Specification and Decomposition in Reinforcement Learning. *ICML* (long oral), pages 2112-2121, 2018.

Toro Icarte, R., Klassen, T., Valenzano, R., McIlraith, S. Teaching Multiple Tasks to an RL Agent using LTL. *AAMAS*, pages 452-461, 2018.

Camacho, A., Baier, J., Muise, C., McIlraith, S. Finite LTL Synthesis as Planning. *ICAPS*, pages 29-38, 2018.

Camacho, A., Baier, J., Muise, C., McIlraith, S. SynKit: LTL Synthesis as a Service. *IJCAI*, pages 5817-5819, 2018.

Klassen, T., McIlraith, S., Levesque, H. Specifying Plausibility Levels for Iterated Belief Change in the Situation Calculus. *KR*, pages 257-266, 2018.

Camacho, A., Bienvenu, M., McIlraith, S. Finite LTL Synthesis with Environment Assumptions and Quality Measures. *KR*, pages 454-463, 2018.

Camacho, A., Baier, J., Muise, C., McIlraith, S. SynKit: Finite LTL Synthesis as a Service. *ICAPS Demo*, 2018. (3 pages)

Camacho, A., Baier, J., Muise, C., McIlraith, S. Synthesizing controllers: On the Correspondence Between LTL Synthesis and Non-Deterministic Planning. *AI*, pages 45-59, 2018.

Toro Icarte, R., Klassen, T., Valenzano, R., McIlraith, S. Advice-Based Exploration in Model-Based Reinforcement Learning. *AI*, pages 72-83, 2018.

Shvo, M. Sohrabi, S., McIlraith, S. An AI Planning-Based Approach to the Multi-Agent Plan Recognition Problem. *AI*, pages 253-258, 2018.

Illanes, L., McIlraith, S. Numeric Planning via Abstraction and Policy Guided Search *IJCAI*, pages 4338-4345, 2017.

Toro Icarte, R., Klassen, T., Valenzano, R., McIlraith, S. Advice-Based Exploration in Model-Based Reinforcement Learning. *RLDM*, pages 199–203, 2017. [*]

Camacho, A., Chen, O., Sanner, S., McIlraith, S. Decision-Making with Non-Markovian Rewards: Guiding search via MDP transformations. *RLDM*, pages 279–283, 2017. [*]

Camacho, A., Chen, O., Sanner, S., McIlraith, S. Non-Markovian Rewards Expressed in LTL: Guiding Search Via Reward Shaping. Camacho, A. *SoCS*, pages 159-160, 2017.

Camacho, A., Triantafillou, E., Muise, C., Baier, J., McIlraith, S. Non-Deterministic Planning with Temporally Extended Goals: LTL over finite and infinite traces. *AAAI*, pages 3716-3724, 2017.

Mombourquette, B., Muise, C., McIlraith, S. Logical Filtering and Smoothing: State Estimation in Partially Observable Domains. *AAAI*, pages 3613–3621, 2017.

Klassen, T., Levesque, H., McIlraith, S. Towards Representing What Readers of Fiction Believe. *13th International Symposium on Commonsense Reasoning*, 2017. (8 pages)

Illanes, L., McIlraith, S. Numeric Planning via Search Space Abstraction (Extended Abstract). *SoCS*, pages 133–134, 2016.

Camacho, A., Muise, C., McIlraith, S. From FOND to Robust Probabilistic Planning: Computing compact policies that bypass avoidable deadends. *ICAPS*, pages 65–69, 2016.

Luo, R., Valenzano, R. A., Li, Y., Beck, J. C., McIlraith, S. Using Metric Temporal Logic to Specify Scheduling Problems. *KR*, pages 581–584, Cape Town, 2016.

Yoon, Y., Robinson, N., Muthusamy, V., McIlraith, S., Jacobsen, H. Planning the transformation of overlays. *Proceedings of the 31st Annual ACM Symposium on Applied Computing (SAC)*, pages 500–507, 2016.

Yoon, Y., Robinson, N., Muthusamy, V., McIlraith, S., Jacobsen, H. Towards Planning the Transformation of Overlays. *Proceedings of the 35th IEEE International Conference on Distributed Computing Systems (ICDCS)*, pp. 782-783, 2015.

Muise, C., Belle, V., Felli, P., McIlraith, S., Miller, T., Pearce, A., Sonenberg, L. Planning Over Multi-Agent Epistemic States: A Classical Planning Approach. *AAAI*, pp. 3327-3334, Austin, Texas, 2015.

Muise, C., Belle, V., and McIlraith, S. Computing Contingent Plans via Fully Observable Non-Deterministic Planning. *AAAI*, pp. 2322-2329, Quebec City, 2014.

Robinson, N., McIlraith, S., and Toman, D. Cost-Based Query Optimization via AI Planning. *AAAI*, pp. 2344-2351, Quebec City, 2014.

Baier, J., Mombourquette, B., and McIlraith, S. Diagnostic Problem Solving via Planning with Ontic and Epistemic Goals. *KR*, Vienna, Austria, 2014. (10 pages)

Muise, C., Belle, V., and McIlraith, S. Non-Deterministic Planning With Conditional Effects *ICAPS*, Portsmouth, USA, 2014. (5 pages)

Muise, C., Beck, J. C., and McIlraith, S. Flexible Execution of Partial Order Plans With Temporal Constraints. *IJCAI*, Beijing, China, 2013. (8 pages)

Davis-Mendelow, Sammy, Baier, J., and McIlraith, S. Assumption-Based Planning: Generating Plans and Explanations under Incomplete Knowledge. *AAAI*, Seattle, USA, 2013. (8 pages)

Muise, C., McIlraith, S., and Beck, J. C. Improved Non-deterministic Planning by Exploiting State Relevance. *ICAPS*, pp. 172-180, Atibaia, Brazil, June, 2012.

Muise, C., McIlraith, S., and Beck, J. C. Optimally Relaxing Partial-Order Plans with MaxSAT. *ICAPS*, pp. 358-362, Atibaia, Brazil, June, 2012.

Muise, C., McIlraith, S., Beck, J.C., and Hsu, E.I. DSHARP: Fast d-DNNF Compilation with sharpSAT. *AI*, pp. 356-361, Toronto, Canada, May, 2012.

Juma, F., Hsu, E.I., and McIlraith, S. Preference-Based Planning with MaxSAT. *AI*, pp. 109-120, Toronto, Canada, May, 2012.

Sohrabi, S., Baier, J., and McIlraith, S. Preferred Explanations: Theory and Generation via Planning. *AAAI*, pp. 261-267, San Francisco, USA, August, 2011.

Muise, C., McIlraith, S., and Beck, J. C. Monitoring the Execution of Partial-Order Plans via Regression. *IJCAI*, pp. 1975-1982, Barcelona, Spain, July, 2011.

Sohrabi, S., and McIlraith, S. Preference-based Web Service Composition: A Middle Ground Between Execution and Search. *ISWC*, pp. 713-729, Shanghai, China, November, 2010.

Liaskos, S., McIlraith, S., Sohrabi, S., and Mylopoulos, J. Integrating Preferences into Goal Models for Requirements Engineering. In Proceedings of the 10th International Requirements Engineering Conference (RE-10), Sydney, Australia, September, 2010.

Sohrabi, S., Baier, J., and McIlraith, S. Diagnosis as Planning Revisited. *KR*, pp. 26-36, Toronto, Canada, May, 2010.

Liaskos, S., McIlraith, S., and Mylopoulos, J. Towards Augmenting Requirements Models with Preferences. In Proceedings of the 24th IEEE/ACM International Conference on Automated Software Engineering (ASE 2009), pp. 565-569, 2009.

Sohrabi, S. and McIlraith, S. Optimizing Web Service Composition while Enforcing Regulations. *ISWC*, pp. 601-617, Chantilly, VA, October, 2009.

Fritz, C. and McIlraith, S. Computing Robust Plans in Continuous Domains. *ICAPS*, pp. 346-349, Thessaloniki, Greece, September 19 – 23, 2009.

Muise, C., McIlraith, S., Baier, J. A., and Reimer, M. Exploiting N-gram Analysis to Predict Operator Sequences. *ICAPS*, Thessaloniki, Greece, September 19 – 23, 2009.

Hsu, E. and McIlraith, S. VARSAT: Integrating Novel Probabilistic Techniques with DPLL Search. *SAT*, pp. 377-390, Swansea, Wales, June 30 – July 3, 2009.

Sohrabi, S., Baier, J., and McIlraith, S. HTN Planning with Preferences. *IJCAI*, pp.1790-1797. Pasadena, California, 2009.

Fritz, C. and McIlraith, S. Generating Optimal Plans in Highly-Dynamic Domains. *UAI*, Montreal, Canada. June 18-21, 2009.

Hsu, E., Muise, C., Beck, J.C., and McIlraith, S. Probabilistically Estimating Backbones and Variable Bias: Experimental Overview. *CP*, 2008.

Fritz, C., Baier, J., and McIlraith, S. ConGolog, Sin Trans: Compiling ConGolog into Basic Action Theories for Planning and Beyond. *KR*, pp. 600-610, Sydney, Australia, September 2008.

Binas, A. and McIlraith, S. Peer-to-peer Query Answering with Inconsistent Knowledge. *KR*, pp. 329-339, Sydney, Australia, September, 2008.

Baier, J., Fritz, C., Bienvenu, M., and McIlraith, S. Beyond Classical Planning: Procedural Control Knowledge and Preferences in State-of-the-Art Planners. *AAAI*, pp 1509-1512, 2008.

Fritz, C. and McIlraith, S. Planning in the Face of Frequent Exogenous Events. *ICAPS Online Poster Proceedings.*, 2008.

Wang, Y., McIlraith, S., Yu, Y., and Mylopoulos, J. An Automated Approach for Monitoring and Diagnosing Requirements. Proceedings of the Twenty-Second IEEE/ACM International Conference on Automated Software Engineering (ASE'07), 10 pages, Atlanta, USA, November, 2007. [Distinguished Paper Award]

Baier, J., Fritz, C. and McIlraith, S. Exploiting Procedural Domain Control Knowledge in State-of-the-Art Planners, *ICAPS*, pp 26–33, Providence, USA, September, 2007.

Fritz, C. and McIlraith, S. Monitoring Plan Optimality During Execution, *ICAPS*, pp 144-151, Providence, USA, September, 2007.

Hsu, E., Kitching, M., Bacchus, F., and McIlraith, S. Using Expectation Maximization to Find Likely Assignments for Solving CSP's. *AAAI*, pp 224–230, Vancouver, Canada, July, 2007.

Baier, J., Bacchus, F., and McIlraith, S. A Heuristic Search Approach to Planning with Temporally Extended Preferences. *IJCAI*, pp 1808-1815, Hyderabad, India, January, 2007.

Sohrabi, S., Prokoshynka. N., and McIlraith, S. Web Service Composition via Generic Procedures and Customizing User Preferences. *ISWC*, pp. 597–611, Athens, USA, November, 2006.

Hsu, E. and McIlraith, S. Characterizing Propagation Methods for Boolean Satisfiability. *SAT*, pp. 325–338, Seattle, USA, August 12 - 15, 2006.

Baier, J. and McIlraith, S. Planning with First-order Temporally Extended Goals Using Heuristic Search. *AAAI*, pp. 788–795, Boston, USA, July, 2006.

Baier, J. and McIlraith, S.. On Planning with Programs that Sense. *KR*, pp. 492–502, Lake District, UK, June, 2006.

Bienvenu, M., Fritz, C. and McIlraith, S. Planning with Qualitative Temporal Preferences. *KR*, pp. 134–144, Lake District, UK, June, 2006.

Fritz, C. and McIlraith, S. Decision-Theoretic Golog with Qualitative Preferences. *KR*, pp. 153–163, Lake District, UK, June, 2006.

Sanner, S. and McIlraith, S. An Ordered Theory Resolution Calculus for Hybrid Reasoning in First-order Extensions of DLs. *KR*, pp. 100–111, Lake District, UK, June, 2006.

Baier, J. and McIlraith, S.. Planning with Temporally Extended Goals using Heuristic Search. *ICAPS*, pp. 342–345, Lake District, UK, June, 2006.

Mandell, D. and McIlraith, S.. Adapting BPEL4WS for the Semantic Web: The Bottom-Up Approach to Web Service Interoperation. *Proceedings of the 2nd International Semantic Web Conference (ISWC2003)*, pp. 227–241, Sanibel Island, FL, USA, October 20-23, 2003.

MacCartney, B., McIlraith, S., Amir, E., Uribe, T.. Practical Partition-Based Theorem Proving for Large Knowledge Bases. *IJCAI*, pp. 89–96, Acapulco, Mexico, August 9-15, 2003.

Lerner, U., Moses, B., Scott, M., McIlraith, S., and Koller, S.. Monitoring a Complex Physical System using a Hybrid Dynamic Bayes Net. *UAI*, pp. 301–310, University of Alberta, Edmonton, August 1-4, 2002.

Narayanan, S. and McIlraith, S.. Simulation, Verification and Automated Composition of Web Services. *Proceedings of the Eleventh International World Wide Web Conference (WWW-2002)*, pp. 77–88, Honolulu, Hawaii, USA, May 7-11, 2002.

McIlraith, S. and Son, T.. Adapting Golog for Composition of Semantic Web Services. *KR*, pp. 482–493, Toulouse, France, April 22-25, 2002.

DAML Services Coalition (alphabetically): A. Ankolekar, M. Burstein, J. Hobbs, O. Lassila, D. Martin, D. McDermott, S. McIlraith, S. Narayanan, M. Paolucci, T. Payne, K. Sycara.. DAML-S: Web Service Description for the Semantic Web. *ISWC*, pp. 348–363, Sardinia, Italy, July 9-12, 2002.

Son, T. Baral, C. and McIlraith, S.. Planning with Different Forms of Domain-Dependent Control Knowledge - An Answer Set Programming Approach. *Proceedings of the 6th International Conference on Logic Programming and Nonmonotonic Reasoning (LPNMR 2001)*, pp. 226-239, Vienna, Austria, September, 2001.

McIlraith, S. and Amir, E.. Theorem Proving with Structured Theories. *IJCAI*, pp. 624–631, Seattle, WA, August, 2001.

DAML Services Coalition (alphabetically): Ankolekar, A., Burstein, M., Hobbs, J., Lassila, O., Martin, D., McIlraith, S., Narayanan, S., Paolucci, M., Payne, T., Sycara, K., Zeng, H.. DAML-S: Semantic Markup for Web Services. *Proceedings of the International Semantic Web Working Symposium (SWWS)*, pp. 411-430, Stanford, CA, July, 2001.

McIlraith, S. and Scherl, R.. What Sensing Tells Us: Towards a Formal Theory of Testing for Dynamical Systems. *AAAI*, pp. 483–490, Orlando, FL, August, 2000.

Amir, E. and McIlraith, S.. Partition-Based Logical Reasoning. *KR*, pp. 389–400, Breckenridge, CO, April, 2000.

Baral, C., McIlraith, S., Tran, S.. Formulating Diagnostic Problem Solving Using an Action Language with Narratives and Sensing. *KR*, pp. 311–322, Breckenridge, CO, April, 2000.

McIlraith, S.. Explanatory Diagnosis: Conjecturing Actions to Explain Observations. *KR*, pp. 167–177, Trento, Italy, June, 1998.

McIlraith, S.. Representing Actions and State Constraints in Model-Based Diagnosis. *AAAI*, pp. 43–49, Providence, RI, July, 1997.

McIlraith, S.. Generating Tests using Abduction. *KR*, pp. 449-460, Bonn, Germany, 1994.

McIlraith, S.. Qualitative Modeling: Application of a Mechanism for Interpreting Graphical Data. *Proceedings of the Seventh Biennial Conference of the Canadian Society for Computational Studies of Intelligence (CSCSI'88)*, pp. 255-261, Edmonton, Alberta, Canada, June 6-10, 1988.

REFEREED WORKSHOP AND SYMPOSIUM PUBLICATIONS

CommonSense:	*International Symposium on Logical Formalizations of Commonsense Reasoning (biennial)*
Commonsense:	*International Symposium on Commonsense Reasoning (biennial)*
DX:	*International Workshop on Principles of Diagnosis*
ICAPS:	*International Conference on Automated Planning and Scheduling*
NAACL:	*Annual Conference of the North American Chapter of the Association for Computational Linguistics*
NRAC:	*International Workshop on Nonmonotonic Reasoning, Action and Change (biennial)*
NeurIPS:	*Conference on Neural Information Processing Systems*
SARA:	*Symposium on Abstraction, Reformulation and Approximation (biennial)*

Chen, D.Z., Zenn, J., Cinquin, T., McIlraith, S.A. Language Models For Generalised PDDL Planning: Synthesising Sound and Programmatic Policies. In *PRL*: Workshop on Programmatic Reinforcement Learning @ RLC 2025, and *EWRL*: 18th European Workshop on Reinforcement Learning (EWRL), 2025.

Chen, D.Z., Hofmann, T., Klassen, T.Q., and McIlraith, S.A. MOOSE: Satisficing and Optimal Generalised Planning via Goal Regression. In *TKR*: 1st International Workshop on Trends in Knowledge Representation and Reasoning @ IJCAI 2025, 2025.

Alamdari, P.A., Klassen, T.Q., and McIlraith, S.A. Auditing, Monitoring, and Intervention for Compliance of Advanced AI Systems. In *R2FM*: Workshop on Reliable and Responsible Foundation Models @ ICML 2025.

McKinney, L., Thudi, A., Bae, J., Kheirkhah, T.R., Papernot, N., McIlraith, S.A., Grosse, R.B. Gauss-Newton Unlearning for the LLM Era. In *MUGen*: Workshop on Machine Unlearning in Generative AI. @ ICML 2025, 2025.

Schiavone, N., Cohen, E., McIlraith, S.A. It's Rational for AI Agents to Procrastinate. In Workshop on Multi-Agent Systmes in the Era of Foundation Models: Opportunities, Challenges and Futures @ ICML 2025

Lifshitz, S., McIlraith, S.A., and Du, Y. Multi-Agent Verification: Scaling Test-Time Compute with Multiple Verifiers (Abridged Report). In Workshops VerifAI, MCDC, SSI-FM, and Reasoning and Planning for LLMs @ ICLR 2025, (arXiv 2502.20379.) 2025.

Klassen, T. Q., Alamdari, P. A., and McIlraith, S.A. Pluralistic Alignment Over Time. In Pluralistic Alignment @ NeurIPS 2024 Workshop, 2024.

Alamdari, P. A., Klassen, T. Q., and McIlraith, S.A. Being Considerate as a Pathway Towards Pluralistic Alignment for Agentic AI. In Pluralistic Alignment @ NeurIPS 2024 Workshop, 2024.

Alamdari, P. A., Klassen, T. Q., Creager, E., and McIlraith, S.A. Remembering to Be Fair: On Non-Markovian Fairness in Sequential Decision Making.Reinforcement Learning Safety Workshop @ Reinforcement Learning Conference (RLC), 2024. (Reprint of ICML 2024 paper.)

Lifshitz, S., Paster, K., Chan, H., Ba, J., and McIlraith, S. A Generative Model for Text Control in Minecraft. In ICML 2023 Workshop on Interactive Learning with Implicit Human Feedback (ILHF), 2023.

Lifshitz, S., Paster, K., Chan, H., Ba, J., and McIlraith, S. Generative Model for Text Control in Minecraft (Abridged Version). In ICML 2023 Workshop on Structured Probabilistic Inference & Generative Modeling, 2023.

Klassen, T. Q., Muise, C., and McIlraith, S. A. Towards Human-Aware AI via Planning with Epistemic Preferences. In ICAPS 2023 Workshop on Human-Aware Explainable Planning (HAXP), 2023.

Klassen, T., Alizadeh Alamdari, P., and McIlraith, S. Epistemic Side Effects & Avoiding Them (Sometimes) In NeurIPS 2022 ML Safety Workshop, 2022.

Paster, K., Pitis, S., McIlraith, S. A., and Ba, J. Return Augmentation gives Supervised RL Temporal Compositionality, In NeurIPS 2022 Deep Reinforcement Learning Workshop, 2022. *This paper also appeared in the NeurIPS 2022 Foundation Models for Decision Making Workshop.*

Li, A., Chen, Z., Vaezipoor, P., Klassen, T., Toro Icarte, R., McIlraith, S. Noisy Symbolic Abstractions for Deep RL: A case study with Reward Machines, In NeurIPS 2022 Deep Reinforcement Learning Workshop, 2022.

Paster, K., McIlraith, S. A., and Ba, J. You Can't Count on Luck: Why Decision Transformers Fail in Stochastic Environments. In ICML 2022 Decision Awareness in Reinforcement Learning Workshop, 2022.

Vaezipoor, P. Li, A., Toro Icarte, R., McIlraith, S. Exploring Long-Horizon Reasoning with Deep RL in Combinatorially Hard Task. In ICML 2022 Decision Awareness in Reinforcement Learning Workshop, 2022.

Tuli, M., Li, A., Vaezipoor, P., Klassen, T. Q., Sanner, S., and McIlraith, S. A. Instruction Following in Text-based Games. In NAACL 2022 Wordplay: When Language Meets Games Workshop, 2022.

Klassen, T. Q., McIlraith, S. A., and Muise, C. An AI Safety Threat from Learned Planning Models. In ICAPS 2022 Workshop on Reliable Data-Driven Planning and Scheduling (RDDPS), 2022.

Paster, K., McKinney, L. E, McIlraith, S. A., and Ba, J. BLAST: Latent Dynamics Models from Bootstrapping. In NeurIPS 2021 Deep RL Workshop, 2021.

Alizadeh Alamdari, P., Klassen, T. Q., Toro Icarte, R., and McIlraith, S. A. Avoiding Negative Side Effects by Considering Others. In NeurIPS 2021 Workshop on Safe and Robust Control of Uncertain Systems, 2021. (More mature version of work to appear at AAMAS 2022.)

Klassen, T. Q., and McIlraith, S. A. Planning to Avoid Side Effects (Preliminary Report). In IJCAI Workshop on Robust and Reliable Autonomy in the Wild (R2AW), 2021. (More mature version of work appeared at AAAI 2022.)

Shvo, M., Klassen, T. Q., and McIlraith, S. A. Explaining the Plans of Agents via Theory of Mind. In ICAPS 2021 Workshop on Explainable AI Planning (XAIP), 2021. (More mature version of work to appear at ICAPS 2022.)

Shvo, M., Klassen, T. Q., and McIlraith, S. A. Towards the Role of Theory of Mind in Explanation. In Explainable, Transparent Autonomous Agents and Multi-Agent Systems - Second International Workshop, EXTRAAMAS 2020, volume 12175, of Lecture Notes in Computer Science, pages 75-93, 2020. Springer

Paster, K., McIlraith, S., Ba, J., Planning from Pixels using Inverse Dynamics Models. In Deep Learning Workshop, at NeurIPS, 2020. *(Longer paper appeared at ICLR-21.)*

Toro Icarte, R., Valenzano, R., Klassen, T., Christoffersen, P., Farahmand, A., McIlraith, S. (2020). The Act of Remembering: A Study in Partially Observable RL – RWRL Report. In: Proceedings of the Challenges of Real World Reinforcement Learning Workshop at NeurIPS 2020.

Christoffersen, P. J., Li, A. C., Toro Icarte, R., and McIlraith, S. A. Learning Symbolic Representations for Reinforcement Learning of Non-Markovian Behavior. In 4th Knowledge Representation and Reasoning Meets Machine Learning Workshop (KR2ML 2020), at NeurIPS, 2020.

Shvo, M., Li, A. C., Toro Icarte, R., and McIlraith, S. A. Interpretable Sequence Classification via Discrete Optimization (Abridged Report). In 4th Knowledge Representation and Reasoning Meets Machine Learning Workshop (KR2ML 2020), at NeurIPS, 2020. *(Longer paper appeared at AAAI-21.)*

Illanes, L., Yan, X., Icarte, R. T., and McIlraith, S. A Symbolic plans as high-level instructions for reinforcement learning (Abridged). In 4th Knowledge Representation and Reasoning Meets Machine Learning Workshop (KR2ML 2020), at NeurIPS, 2020. (Best Paper Award) *(Full paper appeared at ICAPS 2020.)*

Shvo, M., Klassen, T., McIlraith, S. Towards the Role of Theory of Mind in Explanation. In International Workshop on Explainable, Transparent Autonomous Agents and Multi-Agent Systems, 2020. Springer

Camacho, A., McIlraith, S. Towards Neural-Guided Program Synthesis of Linear Temporal Logic Specifications. In Workshop on Knowledge Representation and Reasoning Meets Machine Learning (KR2ML) at NeurIPS, 2019.

Camacho, A., Toro Icarte, R., Klassen, T., Valenzano, R., McIlraith, S. LTL and Beyond: Formal Languages for Reward Function Specification in Reinforcement Learning. In Proceedings of the Knowledge Representation and Reasoning Meets Machine Learning workshop at NeurIPS, 2019. *(Reprint of IJCAI 2019 paper.)*

Toro Icarte, R., Waldie, E., Klassen, T., Valenzano, R., Castro, M., McIlraith, S. Learnng Reward Machines for Partially Observable Reinforcement Learning. In Proceedings of the Optimization Foundations of Reinforcement Learning Workshop at NeurIPS, 2019. *(Reprint of NeurIPS 2019 paper.)*

Toro Icarte, R., Illanes, L., Castro, M. P., Cire, A., McIlraith, S., and Beck, J C. Training binarized neural networks using MIP and CP. In Proceedings of the Workshop on Machine Learning with Guarantees at NeurIPS, 2019. *(Reprint of CP 2019 paper.)*

Illanes, L., Yan, X., Toro Icarte, R., McIlraith, S., Leveraging Symbolic Planning Models in Hierarchical Reinforcement Learning. In Proceedings of the Knowledge Representation and Reasoning Meets Machine Learning workshop at NeurIPS, 2019.

Shvo, M., McIlraith, S. Towards Empathetic Planning. *IJCAI 2019 Workshop on Humanizing Artificial Intelligence*, 2019.

Camacho, A., Bienvenu, M., McIlraith, S. Finite LTL Synthesis with Environment Assumptions (Abridged Report). *IJCAI 2019 Workshop on Strategic Reasoning*, 2019. (Extended Version previously published at KR 2018).

Camacho, A., Toro Icarte, R., Klassen, T., Valenzano, R., McIlraith, S. LTL and Beyond: Formal Languages for Reward Function Specification in Reinforcement Learning *IJCAI 2019 Workshop on Declarative Learning Based Programming*, 2019. (Also appears in main conference.)

Toro Icarte, R., Klassen, T., Valenzano, R., McIlraith, S. Teaching Multiple Tasks to an RL Agent using LTL. *NeurIPS (formerly NIPS) 2018 Workshop on Learning by Instruction*, 2018. (Previously published at AAMAS 2018.)

Toro Icarte, R., Klassen, T., Valenzano, R., McIlraith, S. Advice-Based Exploration in Model-Based Reinforcement Learning. *NeurIPS (formerly NIPS) 2018 Workshop on Learning by Instruction*, 2018. (Previously published at AI 2018.)

Camacho, A., Chen, O., Sanner, S., McIlraith, S. Non-Markovian Rewards Expressed in LTL: Guiding Search Via Reward Shaping (Extended Version). *ICML/IJCAI/AAMAS workshop on Goal Specification for Reinforcement Learning*, 2018.

Camacho, A., Baier, J., Muise, C., McIlraith, S. Bridging the Gap Between LTL Synthesis and Automated Planning. *Proceedings of the ICAPS 2018 Workshop on Generalized Planning (GenPlan'17)*, 2017. (10 pages)

Illanes, L., McIlraith, S. Numeric Planning via Abstraction and Policy Guided Search. *Proceedings of the ICAPS 2018 Workshop on Generalized Planning (GenPlan'17)*, 2017. (Published at IJCAI 2018) (8 pages)

Shvo, M. Sohrabi, S., McIlraith, S. An AI Planning-Based Approach to the Multi-Agent Plan Recognition Problem (Preliminary Report). *Proceedings of the AAAI 2017 Workshop on Plan, Activity, and Intent Recognition (PAIR)*, 2017. (9 pages)

Mombourquette, B., Muise, C., McIlraith, S. Belief State Estimation for Planning via Approximate Logical Filtering and Smoothing. *Proceedings of the IJCAI Workshop on Knowledge-based Techniques for Problem Solving and Reasoning, (KnowProS)*, 2016. (9 pages)

Illanes, L., McIlraith, S. Numeric Planning via Search Space Abstraction. *Proceedings of the IJCAI Workshop on Knowledge-based Techniques for Problem Solving and Reasoning, (KnowProS)* 2016. (9 pages)

Camacho, A., McIlraith, S. Strong-Cyclic Planning when Fairness is Not a Valid Assumption. *Proceedings of the IJCAI Workshop on Knowledge-based techniques for problem solving and reasoning (KnowProS)*, 2016. (9 pages)

Camacho, A., Triantafillou, E., Muise, C., Baier, J., McIlraith, S. LTL Synthesis for Non-Deterministic Systems on Finite and Infinite Traces. *Proceedings of the ICAPS Workshop on Domain Independent Heuristics for Planning (HSDIP)*, London, UK, 2016. (9 pages)

Muise, C., McIlraith, S., Beck, J. C., and Hsu, E. DSHARP: Fast d-DNNF Compilation with sharpSAT (Amended Version). *AAAI-16 Workshop on Beyond NP* Phoenix, Arizona, 2016. (8 pages)

Muise, C., Belle, V., Felli, P., McIlraith, S., Miller, T., Pearce, A., Sonenberg,L. Planning Over Multi-Agent Epistemic States: A Classical Planning Approach (Amended Version). *Proceedings of the ICAPS Workshop on Distributed and Multi-Agent Planning (DMAP)*, Jerusalem, 2015. (8 pages)

Camacho, A., Muise, C., Ganeshen, A., McIlraith, S. From FOND to Probabilistic Planning: Guiding search for quality policies. *Proceedings of the ICAPS Workshop on Domain Independent Heuristics for Planning (HSDIP)*, Jerusalem, 2015. (9 pages)

Triantafillou, E., Baier, J., McIlraith, S., A Unifying Framework for Planning with LTL and Regular Expressions. *Proceedings of the ICAPS Workshop on Model Checking and Automated Planning (MOCHAP)*, Jerusalem, 2015. (9 pages)

Klassen, T., McIlraith, S., Levesque, H., Towards Tractable Inference for Resource-Bounded Agents. *Commonsense*. Stanford, 2015. (6 pages)

Hahmann, T., McIlraith, S. Towards Ontologies in Variation. *Proceedings of the AAAI Spring Symosium on Knowledge Representation and Reasoning (KRR)*. Stanford, 2015. (4 pages)

Muise, C., Belle, V., and McIlraith, S. Computing Contingent Plans via Fully Observable Non-Deterministic Planning. *Proceedings of the ICAPS Workshop on Models and Paradigms for Planning Under Uncertainty (MPPU)*, New Hampshire, 2014. (9 pages)

Baier, J., Mombourquette, B., and McIlraith, S. Diagnostic Problem Solving via Planning with Ontic and Epistemic Goals (Abridged Version). *Proceedings of the ICAPS Workshop on Models and Paradigms for Planning Under Uncertainty (MPPU)*, New Hampshire, 2014. (9 pages)

Gueroussova, M., Polleres, A., and McIlraith, S. SPARQL with Qualitative and Quantitative Preferences. *Proceedings of the 2nd International Workshop on Ordering and Reasoning (OrdRing)* (at ISWC), Sydney, Australia, October, 2013.

Robinson, N., McIlraith, S., and Toman, D. Query Optimization Revisited: An AI Planning Perspective. *Proceedings of the Scheduling and Planning Applications Workshop (SPARK)* (at ICAPS), Rome, Italy, June, 2013.

Muise, C., Beck, J. C., and McIlraith, S. Flexible Execution of Partial Order Plans With Temporal Constraints. *Proceedings of the Workshop on Planning in Continuous Domains* (at ICAPS), Rome, Italy, June, 2013. To appear.

Davis-Mendelow, S., Baier, J., and McIlraith, S. Making Reasonable Assumptions to Plan with Incomplete Information. *Proceedings of the Workshop on Heuristics for Domain Independent Planning* (at ICAPS), Atibaia, Brazil, June, 2012.

Davis-Mendelow, S., Baier, J., and McIlraith, S. Making Reasonable Assumptions to Plan with Incomplete Information (Abridged Report). *Proceedings of the Workshop on Classical Planning for Problem Solving* (at AAAI), Toronto, Canada, July, 2012.

Yoon, Y., Robinson, N., Muthusamy, V., Jacobsen, A., and McIlraith, S. Planned Reconfiguration of Distributed Messaging Systems. *Proceedings of the Workshop on Classical Planning for Problem Solving* (at AAAI), Toronto, Canada, July, 2012.

Muise, C., McIlraith, S., and Beck, J.C. Monitoring the Execution of Partial-Order Plans via Regression *Proceedings of the Workshop on Generalized Planning (GenPlan11)* (at AAAI). San Francisco, USA, July, 2011.

Juma, F., Hsu, E. I., and McIlraith, S. Exploiting MaxSAT for Preference-Based Planning. *Proceedings of the Workshop on Constraint Satisfaction Techniques for Planning and Scheduling Problems (COPLAS)* (at ICAPS). Freiburg, Germany, June, 2011.

Razavi, N., Farzan, A., and McIlraith, S. Predicting Atomicity Violations in Concurrent Programs via Planning. *Proceedings of the Workshop on Verification and Validation of Planning and Scheduling Systems (VVPS)* (at ICAPS). Freiburg, Germany, June, 2011.

Muise, C., McIlraith, S., and Beck, J. C. Optimization of Partial-Order Plans via MAXSAT. *Proceedings of the Workshop on Constraint Satisfaction Techniques for Planning and Scheduling Problems (COPLAS)* (at ICAPS). Freiburg, Germany, June, 2011.

de la Rosa, T. and McIlraith, S. Learning Domain Control Knowledge for TLPlan and Beyond. *Proceedings of the Workshop on Planning and Learning (PAL)* (at ICAPS). Freiburg, Germany, June, 2011.

Sohrabi, S., Baier, J., and McIlraith, S. Diagnosis as Planning Revisited: Abridged Report. *DX*, Portland, OR, 2010.

Hsu, E.I. and McIlraith, S. Computing Equivalent Transformations for Combinatorial Optimization by Branch-and-Bound Search. *Proceedings of Third Annual Symposium on Combinatorial Search (SoCS '10)* Stone Mountain, GA. July, 2010.

Muise, C., McIlraith, S., Beck, J.C., and Hsu, E.I. Fast d-DNNF Compilation with sharpSAT. *Proceedings of the Workshop on Abstraction, Reformulation, and Approximation (WARA-10)* (at AAAI), Atlanta, GA, USA, July, 2010.

Wang, L., Baier, J., and McIlraith, S. Viewing Landmarks as Temporally Extended Goals. *Proceedings of the Workshop on Heuristics for Domain Independent Planning* (at ICAPS). Thessaloniki, Greece, September 19 – 23, 2009.

Albarghouthi, A., Baier, J., and McIlraith, S. On the Use of Planning Technology for Verification. *Proceedings of the Workshop on Heuristics for Domain Independent Planning* (at ICAPS). Thessaloniki, Greece, September 19 – 23, 2009.

Muise, C., McIlraith, S., Baier, J., and Reimer, M. Exploiting N-gram Analysis to Predict Operator Sequences. *Proceedings of the Workshop on Planning and Learning* (at ICAPS). Thessaloniki, Greece, September 19 – 23, 2009.

Sohrabi, S., Baier, J., and McIlraith, S. HTN Planning with Preferences *CommonSense*, Toronto, Canada, May, 2010.

Sohrabi, S., Baier, J., and McIlraith, S. HTN Planning with Quantitative Preferences via Heuristic Search. *Proceedings Oversubscribed Planning and and Scheduling Workshop* (at ICAPS), Sydney, Australia, September, 2008.

Sohrabi, S. and McIlraith, S. On Planning with Preferences in HTN. *Proceedings Fourth Multidisciplinary Workshop on Advances in Preference Handling (MPref08)* (at AAAI), 103–109, 2008.

Sohrabi, S., and McIlraith, S. On Planning with Preferences in HTN (Extended Version). *Proceedings Twelfth International Workshop on Non-Monotonic Reasoning (NMR08)* (at KR), 2008.

Hsu, E., Muise, C., Beck, J. C., and McIlraith, S. Applying Probabilistic Inference to Heuristic Search by Estimating Variable Bias, *Proceedings of the 1st International Symposium on Search Techniques in Artificial Intelligence and Robotics* (at AAAI), 2008.

Fritz, C. and McIlraith, S. Planning in the Face of Frequent Exogenous Events. *Proceedings of the 1st International Symposium on Search Techniques in Artificial Intelligence and Robotics* (at AAAI), 2008.

Fritz, C. and McIlraith, S. Monitoring Policy Execution, *Proceedings of the 3rd Workshop on Planning and Plan Execution for Real-World Systems* (at ICAPS), September, 2007.

Binas, A. and McIlraith, S. Exploiting Prefereces over Information Sources to Efficiently Resolve Inconsistencies in Peer-to-peer Query Answering, *Proceedings of the 2007 AAAI Workshop on Preference Handling for Artificial Intelligence* (at AAAI), pp 15–22, July, 2007.

Fritz, C. and McIlraith, S. Monitoring Plan Optimality during Execution: Theory and Implementation. *DX*, pp 298–305, May, 2007.

Baier, J. and McIlraith, S. On Domain-Independent Heuristics for Planning with Qualitative Preferences. *NRAC* (at IJCAI) January, 2007. A summarized version also appeared in *Commonsense*, March, 2007.

Zeng, H. and McIlraith, S. Experimental Results on the Satisfiable Core in Random 3-SAT. *Proceedings of the 9th International Symposium on Artificial Intelligence and Mathematics*, Fort Lauderdale, Florida, January 4-6, 2006.

Bienvenu, M. and McIlraith, S. Qualitative Dynamical Preferences in the Situation Calculus. *Proceedings of the IJCAI-05 Multidisciplinary Workshop on Advances in Preference Handling*, pp. 13–18, July 31, 2005.

Baier, J. and McIlraith, S. Planning with Programs that Sense. *NRAC* (at IJCAI), pp. 7–14, August, 2005.

Fritz, C. and McIlraith, S. Compiling Qualitative Preferences into Decision-Theoretic GOLOG, *NRAC* (at IJCAI), pp. 45–52, August, 2005.

Bienvenu, M. and McIlraith, S. Specifying and Generating Preferred Plans. *CommonSense*, pp. 25–31, Corfu, Greece, May 22-24, 2005.

Martin, D., Paolucci, M., McIlraith, S., Burstein, M., McDermott, D., McGuinness, D., Parsia, B., Payne, T., Sabou, M., Solanki, M., Srinivasan, N., and Sycara, K. Bringing Semantics to Web Services: The OWL-S Approach. Proceedings of the First International Workshop on Semantic Web Services and Web Process Composition (SWSWPC 2004), pp. 26–42, San Diego, CA, USA, July 6-9, 2004.

Mandell, D. and McIlraith, S. A Bottom-Up Approach to Automating Web Service Discovery, Customization, and Semantic Translation. *Proceedings of the Twelfth International World Wide Web Conference, Workshop on E-Services and the Semantic Web (ESSW'03)*, Budapest, 2003.

McIlraith, S. and Fadel, R. Planning with Complex Actions. *Proceedings of the Ninth International Workshop on Non-Monotonic Reasoning (NMR2002)*, pp. 356-364, April, 2002.

Amir, E. and McIlraith, S. Solving Satisfiability using Decomposition and the Most Constrained Subproblem (Preliminary Report). *Proceedings of LICS 2001 Workshop on Theory and Applications of Satisfiability Testing (SAT 2001)*, Electronic Notes in Discrete Mathematics, Volume 9, Elsevier Publishing, edited by Henry Kautz and Bart Selman, Boston, MA, June 14-15, 2001.

Amir, E. and McIlraith, S. Theorem Proving with Structured Theories (Preliminary Report). *Proceedings of LICS 2001 Workshop on Theory and Applications of Satisfiability Testing (SAT 2001)*, Electronic Notes in Discrete Mathematics, Volume 9, Elsevier Publishing, edited by Henry Kautz and Bart Selman, Boston, MA, June 14-15, 2001.

McIlraith, S. and Son, T. Adapting Golog for Programming the Semantic Web. *CommonSense*, pp. 195–202, New York, NY, May, 2001. Reprinted in *NRAC*.

McIlraith, S., Son, T. and Zeng, H. Mobilizing the Web with DAML-Enabled Web Services. *Proceedings of The Second International Workshop on the Semantic Web (SemWeb'2001)*, Hong Kong, China, May, 2001.

Son, T., Baral, C. and McIlraith, S. Extending Answer Set Planning with Sequence, Conditional, Loop, Non-Deterministic Choice, and Procedure Constructs. *Proceedings of the AAAI Spring Symposium on Answer Set Programming: Towards Efficient and Scalable Knowledge Representation and Reasoning*, pp. 202-209, Stanford, CA, March, 2001.

Amir, E. and McIlraith, S. Improving the Efficiency of Reasoning Through Structure-Based Reformulation. *SARA*, Volume 1864, Lecture Notes in Artificial Intelligence, Springer-Verlag, pp. 247–259, Horseshoe Bay, TX, USA, July 26-29, 2000.

McIlraith, S. Diagnosing Hybrid Systems: A Bayesian Model Selection Approach. *DX*, pp. 140–146, Morelia, Mexico, June, 2000.

McIlraith, S., Biswas, G., Clancy, D., Gupta, V. Hybrid Systems Diagnosis. In Krogh B. and Lynch N. (Eds.), *Hybrid Systems: Computation and Control, Third International Workshop (HSCC'2000)*, Lecture Notes in Computer Science 1790:282–295, Springer-Verlag, Pittsburgh, PA, April, 2000.

McIlraith, S. Modeling and Programming Devices Agents. In Rash, J.L., Rouff, C.A., Truszkowski, W., Gordon, D., and Hinchey M.G. (Eds.), *Formal Approaches to Agent-Based Systems, First International NASA Goddard Workshop*, Lecture Notes in Artificial Intelligence 1871:63–77, Springer-Verlag, Greenbelt, MD, April, 2000.

McIlraith, S. Model-Based Programming using Golog and the Situation Calculus. *DX*, pp. 184–192, Loch Awe, Scotland, June, 1999.

McIlraith, S., Biswas, G., Clancy, D., Gupta, V. Towards Diagnosing Hybrid Systems. *DX*, pp. 193-202, Loch Awe, Scotland, June, 1999.

McIlraith, S., Biswas, G., Clancy, D., Gupta, V. Towards Diagnosing Hybrid Systems (Preliminary Report). *Working Notes of the AAAI 1999 Spring Symposium on Hybrid Systems and AI*, pp. 128-135, Stanford, USA, March, 1999.

McIlraith, S. Towards Exploiting Generic Procedures in Model-Based Reasoning. *DX*, pp. 217–224, Cape Cod, MA, May, 1998.

Choueiry, B., McIlraith, S., Iwasaki, Y., Loeser, T., Neller, T., Engelmore, R., and Fikes, R. Thoughts on a Practical Theory of Reformulation for Reasoning About Physical Systems. *SARA*, pp. 25-36, Pacific Grove, CA, 1998.

McIlraith, S. A Closed-Form Solution to the Ramification Problem (Sometimes). *NRAC*, pp. 103–126, Nagoya, Japan, August, 1997.

McIlraith, S. Explanatory Diagnosis: Conjecturing actions to explain observations. *DX*, pp. 69–78, Mont-Saint-Michel, France, September,1997.

McIlraith, S. SD + Actions: New Representation Problems for Model-Based Diagnosis. *DX*, pp. 157–166, Val Morin, PQ, October, 1996.

McIlraith, S. Towards a Theory of Diagnosis, Testing and Repair. *DX*, pp. 185–192, New Paltz, USA, October, 1994.

McIlraith, S. Generating Tests using Abduction. *DX*, pp. 223-235, Aberystwyth, Wales, September, 1993.

McIlraith, S. Diagnosis as Refutation. *DX*, pp. 232–244, Orcas Island, USA, October, 1992.

McIlraith, S. and Reiter, R. On Experiments for Hypothetical Reasoning. *DX*, pp. 143–152, Milano, Italy, October, 1991. Also appeared as Technical Report RT/DI/91-10-7, Dipartimento di Informatica, Universita di Torino, Torino, Italy, 1991.

LIGHTLY REFEREED WORKSHOP AND SYMPOSIUM PUBLICATIONS

Baier, J., Hussell, J., Bacchus, F., and McIlraith, S. Planning with Temporally Extended Preferences by Heuristic Search. *Proceedings of the ICAPS06 Workshop on Planning with Prefe rences*, pp. 7–10, Lake District, UK, June 2006. A version of this paper also appeared in the Fifth International Planning Competition (IPC-5) Booklet.

Gruninger, M., Hull, R., and McIlraith, S. A First-Order Ontology for Semantic Web Services. (Position Paper), 5 pages. *W3C Workshop on Frameworks for Semantics in Web Services*, Digital Enterprise Research Institute (DERI), Innsbruck, Austria, June 9-10, 2005

Berardi, D., Gruninger, M., Hull, R., McIlraith, S. Towards a First-Order Ontology for Semantic Web Services. (Position Paper), 5 pages. *W3C Workshop on Constraints and Capabilities for Web Services*, Redwood Shore, CA, October, 2005.

Sycara, K., Martin, D., McGuinness, D., McIlraith, S., Paolucci, M. OWL-S Technology for Representing Constraints and Capabilities of Web Services. (Position Paper) 5 pages.*W3C Workshop on Constraints and Capabilities for Web Services*, Redwood Shore, CA, October, 2005.

Choueiry, B., McIlraith, S., Iwasaki, Y., Loeser, T., Neller, T., Engelmore, R. and Fikes, R. Preliminary Thoughts Towards a Practical Theory of Reformulation for Reasoning About Physical Systems. *Working notes of the Workshop on Qualitative Reasoning (QR'98)*, AAAI Technical Report WS-98-01, pp. 21–31, Cape Cod, MA, May 26-29, 1998.

Carlson, B., Fromherz, M., Gupta, V., Hogg, T. and McIlraith, S. Towards Model-Based Adaptive Control of Systems with Smart Matter (Position Paper). *Working Notes of the 1997 AAAI Fall Symposium on Model-Directed Autonomous Systems*, pp. 9–11, Boston, MA, October, 1997.

McIlraith, S. Incorporating Action into Diagnostic Problem Solving (An abridged report). *Working Notes of the 1995 AAAI Spring Symposium on Extending Theories of Actions: Formal Theory and Practical Applications*, pp. 139–144, Stanford University, CA, USA, March 27-29, 1995.

BOOK CHAPTERS

Baier, J., Fritz, C., and McIlraith, S. Golog-Style Search Control for Planning. In Lakemeyer, G. and McIlraith, S. (Eds) *Knowing, Reasoning, and Action, Essays in Honour of Hector J. Levesque.* pp. 47–66. College Publications, 2011.

Sohrabi, S., Prokoshynka. N., and McIlraith, S. Web Services Composition via the Customization of Golog Programs with User Preferences. In. Borgida, A., Chaudhri, V., Giorgini, P., and Yu, E. (Eds) *Conceptual Modeling: Foundations and Applications, Essays in Honor of John Mylopoulos.* pp. 319–334. Springer-Verlag Publishing, 2009.

McIlraith, S.. Explanatory Diagnosis: Conjecturing Actions to Explain Observations. In Levesque, H. and Pirri, F. (Eds.), *Logical Foundations for Cognitive Agents: Papers in Honour of Ray Reiter*, pp. 155–172, Springer-Verlag, 1999.

Pitassi, T., McIlraith, S., and Brecht, T.. Computability. In *Encyclopedia of Electrical and Electronics Engineering*, John Wiley & Sons, Volume 3, pp. 612-618, 1999.

Brecht, T., McIlraith, S., and Pitassi, T.. Recursion. In *Encyclopedia of Electrical and Electronics Engineering*, John Wiley & Sons, Volume 18, pp. 314-315, 1999.

McIlraith, S. and Reiter, R.. On Tests for Hypothetical Reasoning. In Hamschers, W., de Kleer, J. and Console, L. (Eds.), *Readings in Model-Based Diagnosis*, Morgan Kaufmann, pp. 89–95, 1992.

BOOKS EDITED

McIlraith S., and Weinberger, K. (Eds.) Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence, (AAAI-18), the 30th Innovative Applications of Artificial Intelligence (IAAI-18), and the 8th AAAI Symposium on Educational Advances in Artificial Intelligence (EAAI-18), New Orleans, Louisiana, USA, February 2-7, 2018. AAAI Press 2018.

Brewka, G., Eiter, T., and McIlraith, S. (Eds.), Principles of Knowledge Representation and Reasoning: Proceedings of the Thirteenth International Conference, KR 2012, Rome, Italy, June 10-14, 2012. AAAI Press 2012.

Lakemeyer, G. and McIlraith, S. (Eds.) Knowing, Reasoning, and Acting. Essays in Honour of Hector J. Levesque. Tributes 16. College Publications, July, 2011.

Morgenstern, L. and McIlraith, S. (Eds.) Artificial Intelligence Journal. Special Issue: Papers in Honor of John McCarthy. 175(1), January, 2011.

McIlraith, S., Peppas, P., and Thielscher, M. (Eds.), Journal of Logic and Computation. Special Issue: Logical Formalization of Commonsense Reasoning, 17:5, October, 2007.

McIlraith, S., Plexousakis, D. and van Harmelen, F. (Eds.), The Semantic Web – ISWC2004. Proceedings of the International Semantic Web Conference. Hiroshima, Japan, November 2004 Springer-Verlag, 862 pages, October, 2004.

McIlraith, S. and Plexousakis, D. (Eds.), Journal of Web Semantics. Selected Papers from the International Semantic Web Conference, 2004 - ISWC, 2004 Volume 3, Numbers 2-3, 168 pages, Elsevier Publishing, October, 2005.

Bussler, C., Hull, R., McIlraith, S., Orlowska, M.E., Pernici, B., and Yang, J. (Eds.) Web Services, E-Business, and the Semantic Web, CAiSE 2002 International Workshop (WES 2002) SpringerLNCS, Vol. 2512, 277 pages, Springer-Verlag, May, 2002.

UNREFEREED PUBLICATIONS
MAGAZINE ARTICLES

McIlraith, S. and Martin, D.. Bringing Semantics to Web Services. *IEEE Intelligent Systems*, 18(1):90–93, January/February, 2003.

Biswas, G., McIlraith, S.. Report on the 1999 AAAI Spring Symposium – Hybrid Systems and AI: Modeling, Analysis and Control of Discrete + Continuous Systems. *AI Magazine* 21(2), pp. 79-81, 2000.

MAJOR REPORTS

Heintz, F., Lakemeyer, G., McIlraith, S. Cognitive Robotics (Dagstuhl Seminar 22391). Dagstuhl Reports 12(9): 200-219 (2022)

Baral, C., Bolander, T., van Ditmarsch, H., McIlraith, S. Epistemic Planning (Dagstuhl Seminar 17231). Dagstuhl Reports 7(6): 1-47 (2017)

STANDARDIZATION EFFORTS

Battle, S., Bernstein, A., Boley, H. Grosof, B. Gruninger, M., Hull, R., Kifer, M., Martin, D., McIlraith, S., McGuinness, D., Su, J. and Tabet, S.. Semantic Web Services Language (SWSL). *W3C Members Submission*, September 9, 2005. `http://www.w3.org/Submission/SWSF-SWSL/` (*McIlraith was a key contributor.*)

Co-author and key contributor. Semantic Web Services Framework Releases 1.0, 1.1. 2005. `http://www.daml.org/services/swsf/`.

Martin, D., Burstein, M., Hobbs, J., Lassila, O., McDermott, D., McIlraith, S., Narayanan, S., Paolucci, M., Parsia, B., Payne, T., Sirin, E., Srinivasan, N., Sycara, K.. OWL-S: Semantic Markup for Web Services. *W3C Member Submission*, November, 2004. `http://www.w3.org/Submission/OWL-S/` (*McIlraith was a key contributor.*)
Google Scholar Citation Count: 1344

Co-author and key contributor. DAML-S/OWL-S Web Service Ontology Releases 0.5, 0.6, 0.7, 0.9, 1.0, 1.1. 2001 – 2004. `http://www.daml.org/services/owl-s/`.

TECHNICAL REPORTS

Many of these technical reports are (extended) versions of papers that appeared in conferences.

DCS-UT: Department of Computer Science, University of Toronto
KSL-Stanford: Knowledge Systems Lab, Department of Computer Science, Stanford University

Camacho, A., Chen, O., Sanner, S., McIlraith, S. Decision-Making with Non-Markovian Rewards: Guiding search via automata-Based reward shaping. CSRG-632, *DCS-UT*, June 2017.

Luo, R., Beck, J.C., McIlraith, S. A Timed Temporal Logic for Specifying Scheduling Problems. CSRG-631, *DCS-UT* April 2016.

Camacho, A., Muise, C., McIlraith, S. Appendix: From FOND to Robust Probabilistic Planning. CSRG-630, *DCS-UT* March 2016.

Luo, R., Valenzano, R., Li, Y., Beck, J.C., McIlraith, S. A Timed Temporal Logic for Specifying Scheduling Problems (Extended Report). CSRG-629, *DCS-UT* November 2015

Fritz, C., Baier, J., and McIlraith, S. ConGolog Sin Trans: Compiling ConGolog Action Theories into Basic Actions Theories for Planning and Beyond (extended version). CSRG-576, *DCS-UT* December 2008.

Baier, J., Fritz, C. and McIlraith, S. Exploiting Procedural Domain Control Knowledge in State-of-the-Art Planners (extended version). CSRG-565, *DCS-UT*, August, 2007.

Baier, J. and McIlraith, S. Planning with Propositional Temporally Extended Goals. DCS-TR-537, *DCS-UT*, 2006.

Fritz, F. and McIlraith, S. Compiling Qualitative Preferences into Decision-Theoretic Golog Programs: Extended Version with Proofs. CSRG-522, *DCS-UT*, May, 2005.

Zeng, H. and McIlraith, S. Experimental Results on the Satisfiable Core in Random 3SAT. KSL-06-01, *KSL-Stanford*, 2006.

Baier, J. and McIlraith, S. Planning with Programs that Sense. DCS-TR-523, *DCS-UT*, October, 2005

Zeng, H. and McIlraith, S. The Role of Redundant Clauses in Solving Satisfiability Problems (extended abstract). KSL-05-10, *KSL-Stanford*, 2005.

Mandell, D. and McIlraith, S. Bottom-Up Approach to Automating Web Service Discovery, Customization, and Semantic Translation. KSL-03-11, *KSL-Stanford*, 2003.

MacCartney, B., McIlraith, S., Amir, E. and Uribe, T. Practical Partition-Based Theorem Proving for Large Knowledge Bases. KSL-03-12, *KSL-Stanford*, 2003.

Mandell, D. J. and McIlraith, S. Adapting BPEL4WS for the Semantic Web: The Bottom-Up Approach to Web Service Interoperation. KSL-03-13, *KSL-Stanford*, 2003.

Lerner, U., Moses, B., Scott, M., McIlraith, S. and Koller, D. Monitoring a Complex Physical System using a Hybrid Dynamic Bayes Net. KSL-02-08, *KSL-Stanford*, 2002.

Ankolekar, A., Burstein, M., Hobbs, J., Lassila, O., Martin, D., McDermott, D., McIlraith, S., Narayanan, S. and Paolucci, M., Payne, T. and Sycara, K. DAML-S: Web Service Description for the Semantic Web. KSL-02-14, *KSL-Stanford*, 2002.

McIlraith, S. and Son, T. Adapting Golog for Composition of Semantic Web Services. KSL-02-15, *KSL-Stanford*, 2002.

Narayanan, S. and McIlraith, S. Simulation, Verification and Automated Composition of Web Services. KSL-02-16, *KSL-Stanford*, 2002.

McIlraith, S. and Amir, E. Theorem Proving with Structed Theories (Full Report). KSL-01-04, *KSL-Stanford*, April, 2001.

Son, T., Baral, C. and McIlraith, S. Extending Answer Set Planning with Sequence, Conditional, Loop, Non-Deterministic Choice, and Procedure Constructs. KSL-01-13, *KSL-Stanford*, 2001.

McIlraith, S., Son, T. and Zeng, H. Mobilizing the Web with DAML-Enabled Web Services. KSL-01-14, *KSL-Stanford*, 2001.

McIlraith, S. and Son, T. Adapting Golog for Programming the Semantic Web. KSL-01-15, *KSL-Stanford*, 2001.

Amir, E. and McIlraith, S. Solving Satisfiability using Decomposition and the Most Constrained Subproblem. KSL-01-16, *KSL-Stanford*, 2001.

Ankolekar, A., Burstein, M., Hobbs, J., Lassila, O., Martin, D., McIlraith, S., Narayanan, S., Paolucci, M., Payne, T., Sycara, K., and Zeng, H. DAML-S: Semantic Markup for Web Services. KSL-01-17, *KSL-Stanford*, 2001.

Son, T., Baral, C., and McIlraith, S. Planning with Different Forms of Domain-Dependent Control Knowledge - An Answer Set Programming Approach. KSL-01-18, *KSL-Stanford*, 2001.

Amir, E. and McIlraith, S. Partition-Based Logical Reasoning. KSL-00-02, *KSL-Stanford*, February, 2000.

McIlraith, S., Biswas, G., Clancy, D., and Gupta, V. Towards Diagnosing Hybrid Systems. KSL-99-01, *KSL-Stanford*, February, 1999.

McIlraith, S., Biswas, G., Fromherz, M., Howe, J., Fikes, R., Bobrow, D., Cutkosky, M., Engelmore, R., and Neller, T. Model-Enabled Control of Hybrid Systems. KSL-98-22, *KSL-Stanford*, July, 1998.

McIlraith, S. Explanatory Diagnosis: Conjecturing actions to explain observations. KSL-98-21, *KSL-Stanford*, July, 1998.

McIlraith, S. Towards Exploiting Generic Procedures in Model-Based Computing. KSL-92-20, *KSL-Stanford*, July, 1998.

McIlraith, S. Logic-Based Abductive Inference. KSL-98-19, *KSL-Stanford*, July, 1998.

Choueiry, B., McIlraith, S., Iwasaki, Y., Loeser, T., Neller, T., Engelmore, R., and Fikes, R. Thoughts Towards a Practical Theory of Reformulation for Reasoning about Physical Systems. KSL-98-18, *KSL-Stanford*, June, 1998.

McIlraith, S. Logic-Based Abductive Inference. KSL-98-19, *KSL-Stanford*, July, 1998.

McIlraith, S. Towards Exploiting Generic Procedures in Model-Based Computing. KSL-92-20, *KSL-Stanford*, July, 1998.

McIlraith, S. Explanatory Diagnosis: Conjecturing actions to explain observations. KSL-98-21, *KSL-Stanford*, July, 1998.

McIlraith, S., Biswas, G., Fromherz, M., Howe, J., Fikes, R., Bobrow, D., Cutkosky, M., Engelmore, R. and Neller, T. Model-Enabled Control of Hybrid Systems. KSL-98-22, *KSL-Stanford*, July, 1998.

McIlraith, S. and Reiter, R. On tests for hypothetical reasoning. KR-TR-92-2, *DCS-UT*, 1992.

OTHER REPORTS

User Customization of Caregiving Robots that Support Older Adults with Dementia. Zehui Zhou, under the supervision of Professor. Sheila McIlraith, Dr. Steven Shapiro, and Dr. Richard Valenzano. Review of Undergraduate Computer Science (RUCS) an online publication of the University of Toronto. Fall 2017.

Improved Heuristics for Solving the Pancake Problem. Danniel Yang, under the supervision of Dr. Richard Valenzano and Professor Sheila McIlraith. Review of Undergraduate Computer Science (RUCS) an online publication of the University of Toronto. Fall 2016.

An Intelligent Guidance System for People with Dementia in a Picture Description Task. Hengwei Guo, Oscar Chen, Hubert Lin, under the supervision of Alberto Camacho, Hamidreza Chinae, Frank Rudzicz, and Sheila McIlraith. Review of Undergraduate Computer Science (RUCS) an online publication of the University of Toronto. Fall 2015.

A Unifying Framework for Planning with LTL and Regular Expressions. Eleni Triantafillou under the supervision of Jorge Baier and Sheila McIlraith. Review of Undergraduate Computer Science (RUCS) an online publication of the University of Toronto. Fall 2015.

Reports of the AAAI 2014 Conference Workshops. AI Magazine 36(1): 87-98 (2015) Stefano V. Albrecht, Andréa Motta Salles Barreto, Darius Braziunas, David L. Buckeridge, Heriberto Cuayáitl, Nina Dethlefs, Markus Endres, Amir Massoud Farahmand, Mark Fox, Lutz Frommberger, Sam Ganzfried, Yolanda Gil, Séstien Guillet, Lawrence E. Hunter, Arnav Jhala, Kristian Kersting, George Konidaris, Freddy LééSheila A. McIlraith, Sriraam Natarajan, Zeinab Noorian, David Poole, Ré Ronfard, Alessandro Saffiotti, Arash Shaban-Nejad, Biplav Srivastava, Gerald Tesauro, Rosario Uceda-Sosa, Guy Van den Broeck, Martijn van Otterlo, Byron C. Wallace, Paul Weng, Jenna Wiens, Jie Zhang. Reports of the AAAI 2014 Conference Workshops. AI Magazine 36(1): 87-98 (2015).

McIlraith, S. Situation Calculus: The Last 15 Years (Invited Talk). Proceedings of the 14th International Conference on Principles of Knowledge Representation and Reasoning (KR 2014), 2014. (1 page)

McIlraith, S. Towards a Formal Account of Diagnostic Problem Solving. *PhD Dissertation*, Department of Computer Science, University of Toronto, 1997.

McIlraith, S. and Stevenson, S.. The Academic Environment: Graduate School and Beyond. (A synthesis of the findings of participants at a forum for women in science on April 29, 1993). *Manuscript*, Toronto, ON, September, 1993.

Author of over a dozen proprietary reports relating artificial intelligence work for private companies.

Co-author of almost a dozen contract research proposals on behalf of the Knowledge Systems Laboratory, Stanford University, and more than a dozen contract research proposals on behalf of the Alberta Research Council.

INVITED TALKS

Invited Keynote, RLC 2025 Workshop on Programmatic Reinforcement Learning, Edmonton, August 5, 2025. "Programmatic Reward Models: Exploiting reward function structure to help agents learn, plan, and remember"

Invited Keynote, ICML 2025 Workshop on Programmatic Representations for Agent Learning, Vancouver, July 18, 2025. "Programmatic Reward Models: Exploiting reward function structure to help agents learn, plan, and remember"

Invited Lecture, Grad Visit Day, Department of Computer Science, University of Toronto. March 18, 2025. "How (formal) language can help AI agents learn, plan, and remember"

Invited Keynote, AAAI 2025 Workshop on Generalization in Planning (GenPlan). Philadelphia, USA. March 4, 2025. "Generalization for Planning and RL: A story in three acts"

Invited Keynote, AAAI 2025 Workshop on Theory of Mind. Philadelphia, USA. March 3, 2025. "Purposeful Theory of Mind"

Invited Keynote, the 21st International Conference on Principles of Knowledge Representation and Reasoning (KR 2024), November 7, 2024 (hybrid). "Sheila McIlraith & Murray Shanahan in conversation with Joe Halpern - Great Ideas from KR: drawing on the past to shape the future"

Invited Talk, UK Royal Society Event: Beyond the symbols vs signals debate. London, UK, October 29, 2024. "How (formal) language can help AI agents learn, plan, and remember"

Lightning Talk, Bay Area Alignment Workshop, Santa Cruz, CA, October 24, 2024."Using formal languages to encode reward functions, instructions, preferences, norms, and advice ... and so much more ..."

Invited Talk, Aachen Symposium on Representation, Learning to Act and Plan, Vaals, NL. September 16, 2024. "The signal-symbol nexus for sequential decision making: where KR and planning meet machine learning"

Invited Tutorial, 2nd IEEE Conference on Secure and Trustworthy Machine Learning (SaTML), April 11, 2024. "(Formal) Languages Help AI Agents Learn and Reason"

Lightning Talk, New Orleans Alignment Workshop (NOLA 2023), December 11, 2023. "Epistemic Side Effects"

Keynote, RL-CONFORM Workshop at IROS 2023, the IEEE/RSJ International Conference on Intelligent Robots and Systems, October 1, 2023. "Building Taskable Reinforcement Learning Agents via Formal Languages and Automata"

Invited Talk, 1st International Workshop on Knowledge-Based Compositional Generalization (KBCG) at IJCAI 2023, August 21, 2023. "(Formal) Languages Help AI Agents Learn and Reason"

Keynote, 33rd International Conference on Automated Planning and Scheduling (ICAPS 2023), July 12, 2023. "(Formal) Languages Help AI agents Learn, Plan, and Remember"

Invited Talk, Workshop on Bridging the Gap Between AI Planning and Reinforcement Learning (PRL) at ICAPS 2023, July 9, 2023. "Building Taskable RL Agents Using Advice, Instructions, and AI Planning"

Keynote, Midwest ML Symposium (MMLS 2023), May 16, 2023. "(Formal) Languages Help AI Agents Learn and Reason"

Invited Talk, AAAI 2023 Spring Symposium on the Effectiveness of Temporal Logics over Finite Traces in AI, March 27, 2023. "Automated Planning with Temporal Logics: Techniques, Advances, and Broad Applicability"

Keynote, 37th AAAI Conference on Artificial Intelligence (AAAI 2023), February 11, 2023. "(Formal) languages help AI agents learn and reason"

Invited Talk, Distinguished Lecture Series, University of Waterloo, Department of Computer Science, January 11, 2023. "Reward Machines: Formal Languages and Automata for Reinforcement Learning"

Invited Talk, FBK Research Institute, Trento, Italy. November 17, 2022. "Building Taskable Reinforcement Learning Agents via Formal Languages and Automata"

Invited Talk, IROS 2022 Workshop on Integrating Machine Learning and Automated Reasoning, October 27, 2022. "Building Taskable Robots using Reinforcement Learning"

Invited Talk, Dagstuhl Cognitive Robotics Seminar, September 29, 2022. "Reward Machines: Formal Languages and Automata for Reinforcement Learning"

Keynote, 16th International Conference on Logic Programming and Non-monotonic Reasoning (LPNMR 2022), September 6, 2022. "Reward Machines: Formal Languages and Automata for Reinforcement Learning"

Invited Talk, Learning and Automata (LearnAut 2022) at ICALP 2022, July 4, 2022. "Reward Machines: Formal Languages and Automata for Reinforcement Learning"

Invited Talk, University of Toronto Teaching and Learning Symposium, May 12, 2022. "Embedded Ethics Education Initiative" (joint with Horton and Liu)

Invited Talk, COSYNE Workshop on 'The what, how and when of learning: How do artificial and biological agents approach the problem of learning, and how can approaches from one inform the other?' March 22, 2022. "Can (formal) language help RL agents learn and think?"

Invited Talk, IBM Neuro-Symbolic AI Workshop 2022, January 19, 2022. "Building Taskable Reinforcement Learning Agents"

Invited Talk, Schwartz Reisman Institute for Technology and Society (SRI), December 1, 2021. "Can 'Being Considerate' lead to Safer AI?"

Presentation, Teaching and Learning Communties of Practice Meeting, University of Toronto (with Diane Horton and David Liu), December 2, 2021."The Embedded Ethics Education Initiative"

Keynote, WOLVERINE Workshop at the 19th International Symposium on Automated Technology for Verification and Analysis (ATVA), October 18, 2021. "Formal Languages and Automata for Reward Function Specification and Efficient Reinforcement Learning"

Invited Talk, Workshop on Automated Synthesis at the European Summer School on Logic, Language, and Information (ESSLLI 2021), July 29, 2021. "Formal Languages and Automata for Reward Function Specification and Efficient Reinforcement Learning"

Invited Talk, Microsoft Research (MSR) Montreal, July 27, 2021. "Reward Machines: Formal Languages and Automata for Reward Function Specification and Sample Efficient Reinforcement Learning"

Invited Talk, RSS 2021 Workshop on Declarative and Neurosymbolic Representations in Robot Learning and Control at the 2021 Robotics: Science and Systems (RSS) Conference July 15th, 2021 "Building Human-Taskable Robots that Learn"

Invited Talk, European TAILOR project workshop, July 20, 2021. "Formal Languages and Automata for Reward Function Specification and Efficient Reinforcement Learning"

Invited Talk, CHAI Workshop for the Center for Human-Compatible AI, June 7, 2021. "Common Sense, Common Ground, and Avoiding Side Effects"

Invited Talk, Simons Institute for Theory of Computing, Games and Equilibria in System Design and Analysis, May 7, 2021. "Formal Languages and Automata for Reward Function Specification and Efficient Reinforcement Learning"

Invited Talk, Synthetic Intelligence Forum, Toronto, April 21, 2021. "Building Taskable Reinforcement Learning Agents"

Invited Talk, 3rd Annual Rotman CPA Ontario Centre for Accounting Innovation Research Conference, Rotman School of Management, University of Toronto, February 17, 2021. "Responsible AI: Getting your AI to do the right thing"

Invited Talk, Vector Institute Annual Research Symposium, February 16, 2021. "Building Taskable RL Agents"

Invited Talk, 15th Annual Senior College Symposium, *Ethical Challenges of the 21st Century*, November 30, 2020. "Artificial Intelligence: Promise and Peril for Humanity"

Invited Tutorial, International Conference on Automated Planning and Scheduling (ICAPS 2020), October 22, 2020. "Epistemic Planning"

Invited Talk, Great Explorations Series, University of Toronto Scarborough. October 20, 2020. "AI by the People for the People"

Invited Talk, Women in KR, 6 prominent women researchers in Knowledge Representation and Reasoning sharing their vision on the present and future of the field. 17th International Conference on Principles of Knowledge Representation and Reasoning (KR2020), September 17, 2020.

Invited Talk, AAAI 2020 Workshop on Plan, Activity, and Intent Recognition (PAIR 2020), NYC, New York. February 8, 2020. "Behavior Interpretation: Making sense of people, machines, and the world around us"

Invited Talk, AAAI 2020 Workshop on Generalization in Planning (GenPlan 2020), NYC, New York. February 7, 2020. "Learning and Remembering: Planning with and without instruction"

Keynote, Microsoft Research (MSR) Reinforcement Learning Day 2019 NYC, New York. October 3, 2019. "Reward Machines: Structuring reward function specifications and reducing sample complexity in reinforcement learning"

Keynote, International Conference on Logic Programming (ICLP). Las Cruces, New Mexico. September 25, 2019. "Reward Machines: Structuring reward function specifications and reducing sample complexity in reinforcement learning"

Invited Talk, 2019 Vector AI Master's Summit & Career Fair. Design Exchange, Toronto. September 19, 2019. "Doing the Right Thing: and getting your AI to do the same"

Invited Talk, HER CODE CAMP, Academia Panel, Toronto. September 15, 2019. "Why I Love my Job"

Invited Talk, Taking Responsibility for Responsible AI, Munk Centre, Munk School of Global Affairs and Public Policy, University of Toronto. August 26, 2019. "Doing the Right Thing: and getting your AI to do the same"

Keynote, 4th Multidisciplinary Conference on Reinforcement Learning and Decision Making (RLDM), July 8, 2019. "Reward Machines: Structuring reward function specifications and reducing sample complexity in reinforcement learning"

Invited talk, Hybrid Reasoning for Intelligent Systems Workshop, June 21, 2019. "Reward Machines: Structuring reward function specifications and reducing sample complexity in reinforcement learning"

Invited Talk, Imperial College, London, June 6, 2019. "Reward Machines: Structuring reward function specifications and reducing sample complexity in reinforcement learning"

Invited Talk, CIFAR-UKRI-CNRS Workshop *AI and Society: from principles to practice*, Turing Institute, London, June 4, 2019. "Interpretable, Queriable, Taskable: Towards better understanding our AI"

Invited Talk, Ethics of AI in Context, Centre for Ethics, University of Toronto, Toronto, Canada, March 26, 2019. "Making Good Decisions and Getting AI to Do the Same"

Invited Talk, Vector Institute for Artificial Intelligence, Toronto, Canada, December 14, 2018. "High-Level Reward Function Specification and Structural Decomposition in Reinforcement Learning"

Invited Talk, Samsung AI Research Lab, Toronto, Canada, December 11, 2018. "Synthesizing Plans and Programs via Instructions, Advice, and Observations"

Invited Talk, IBM Symposium on Recent Advances in AI Planning, MIT IBM Cambridge AI Lab, Boston, MA, November 13, 2018. "Plan and Program Synthesis: A new look at some old problems"

Invited Talk, Department of Computer Science, Pontificia Universidad Catolica de Chile, Santiago, Chile, November 8, 2018. "Multi-Agent Epistemic Planning"

Keynote, Jornada Chilenas de Computacion (JCC-2018). Santiago, Chile, November 5, 2018. "Artificial Intelligence: Past, Present, and Future"

Invited Talk, Doctoral Consortium, International Joint Conference on Artificial Intelligence (IJCAI), Stockholm, Sweden, July 13, 2018. "Who do you want to be?"

Keynote, First ICAPS Women's Lunch, International Conference on Automated Planning and Scheduling (ICAPS), Delft, Netherlands, June 28, 2018. "Who do you want to be?"

Keynote, International Symposium on Temporal Representation and Reasoning (TIME 2017). Mons, Belgium, October 16, 2017. "Plan and Program Synthesis: A new look at some old problems"

Invited Lecture, Increasing Diversity in AI Workshop at the 31st AAAI Conference on Artificial Intelligence (AAAI), San Francisco, USA, February 5, 2017. "AI and the Future of Automated Programming"

Invited Lecture, Civic Series Talk, Toronto, Canada, December 14, 2016. "Artificial Intelligence: What is it & where is it taking us?"

Invited Lecture, Girl Geeks Toronto, Toronto, Canada, September 27, 2016. "Artificial Intelligence: the biggest tech trend since... the internet?"

Invited Lecture, Summer School in Cognitive Science, Montreal, Canada, June 29, 2016. "Reasoning to Act: From Logic to Automated Planning"

Keynote, 25th International Joint Conference on Artificial Intelligence (IJCAI). New York, USA, July 12, 2016. "Do as I say and as I do: the future of automated programming"

Invited Lecture, Global Smartweek Toronto IoT & Machine Intelligence Conference. Toronto, Canada, November 4, 2015. "Where is machine intelligence heading in R&D within universities and industry"

Invited Lecture, To be Announced! Synthesis of Epistemic Protocols, Lorentz Center, Leiden, Netherlands, August 18, 2015. "Planning Over Multi-Agent Epistemic States: A Classical Planning Approach"

Distinguished Speaker, Oxford University Departmental and OxWoCS Distinguised Speaker Series, Oxford, UK, February 24, 2015. "Of Programs, Plans, and Automata: 101 Things You Can Do with AI Planning"

Invited Lecture, Automated Planning and Model Checking (Dagstuhl seminar 14482), Wadern, Germany, November 23-28, 2014. "LTL and Procedural Control for Planning and Diagnosis"

Keynote, Great Moments in KR Lecture, 14th International Conference on Principles of Knowledge Representation and Reasoning (KR 2014), July 23, 2014. "Situation Calculus: The last 15 years"

Speaker, CBC Quirks & Quarks, The Question Show, Toronto, June 21, 2014. "Does Artificial Intelligence Exist?"

Invited Lecture, Interoperation in Complex Information Ecosystems (Dagstuhl Seminar 13252), Wadern, Germany, June 16-19, 2013. "A Purposeful View of Data: Getting the Data When you Need it and How you Like it"

Keynote, Canadian Conference on Artificial Intelligence, Regina, 2013, May 28, 2013. "From Services to the Cloud and into Space: Modeling and Customizing Component-Based Systems"

Keynote, Math in Motion...Girls in Gear!, University of Toronto at Scarborough, Toronto, November 3, 2012. "Math in Motion: Cognitive Robotics and the Semantic Web"

Invited Lecture, Twenty-Sixth Conference on Artificial Intelligence (AAAI-12), Toronto, July 26, 2012. "What's Hot in Knowledge Representation and Reasoning"

Invited Lecture, Graduate Student Symposium, Canadian Conference on Artificial Intelligence, Toronto, May 27, 2012. "From Services ... to the Cloud ... and into Space: Modeling and Customizing Component-Based System"

Invited Lecture, Symposium in Honor of Professor Dana Nau, University of Maryland, College Park, Maryland, May 24, 2012. "101 Things You Can Do with Non-Classical Planning"

Invited Lecture, Meeting on Artifact-Centric Computing, IBM T.J. Watson, Hawthorne, New York, March 8, 2011. "Modeling, Composing, Verifying and Optimizing Processes: From Web Services to Business Processes"

Invited Lecture, CASCON2010 Workshop on Automated Service Composition, Toronto, November 2010. "Composing, Customizing, and Optimizing Web Services and Business Processes"

Invited Lecture, Girl Geek Dinner Meeting, Toronto, December 9, 2009. "The Semantic Web and Artificial Intelligence"

Distinguished Lecture, University at Buffalo, The State University of New York, Buffalo, NY, November 20, 2009. "Semantic Web Services"

Keynote, International Symposium on Formalization of Commonsense Reasoning. Toronto, Canada, June 2009. "Diagnosis Revisited"

CTV Space Channel, Terminator Salvation: Behind the Scenes. Television interview on Robotics. May 14, 2009.

Invited Lecture, Third International Compulog/ALP Summer School on Logic Programming and Computational Logic. Las Cruces, NM, July 2008. "Foundations of Semantic Web and Computational Logic"

Joint Keynote, International Conference on Automated Planning and Scheduling (ICAPS2007) and International Conference on Constraint Programming (CP2007). Providence, September 2007. "Automated Web Service Compositions: New (and not so new) Challenges for AI Planning"

2004 IBM CASCON Workshop on Semantic Web Services, Toronto, October 2004. "Semantic Web Services"

Keynote, International Conference on Principles and Practice of Declarative Programming, Verona, Italy, August, 2004. "Towards Declarative Programming for Web Services"

Keynote, 2004 International Workshop on Description Logics - DL2004, Whistler, Canada. "Towards Semantic Web Services: OWL-S and Web Service Composition"

Sony Research Labs, San Jose, October, 2002. "Automated Web Service Composition"

Keynote, International Conference on Automated Software Engineering (ASE2002) Edinburgh, September 2002. "Web Services Meet the Semantic Web"

Workshop on Agent-Oriented Information Systems (AOIS02), CAISE02, Toronto, May 2002. "Web Services Meet the Semantic Web"

York University, Toronto, May 2002. "Planning with Complex Actions."

University of British Columbia, Vancouver, August, 2002. "Automating Web Service Composition."

Stanford University, Stanford, May, 2002. "Automating Web Service Composition."

University of Toronto, Toronto, April, 2002. "Automating Web Service Composition."

Queen's University, Kingston, April, 2002. "Automating Web Service Composition."

University of Waterloo, Waterloo, April, 2002. "Automating Web Service Composition."

Invited Tutorial, XML Web Services One, San Jose, CA, June 2002. "Semantic Web Services"

Bell Labs, Lucent Technologies, Murray Hill, New Jersey, October, 2001. "Semantic Web Services."

New Jersey Institute of Technologies, Newark, New Jersey, October, 2001. "Semantic Web Services."

Stanford Networking Research Center Symposium, September 2001. "DAML-S and Services on the Semantic Web."

European IBROW Meeting, Stanford, CA. July, 2001. "DAML-S and Services on the Semantic Web."

Cisco Systems, Mountainview, CA. July, 2001. "DAML-S and Services on the Semantic Web."

Xerox Palo Alto Research Center (PARC). Palo Alto. June, 2001. "Mobilizing the Web with DAML-Enabled Web Services."

University of Linkoping, Linkoping. Sweden. June, 2001. "Mobilizing the Web with DAML-Enabled Web Services."

University of Linkoping, Linkoping. Sweden. June, 2001. "Model-Based Execution Monitoring." Opponent presentation as part of Ph.D. doctoral dissertation for Marcus Bjäreland.

National Research Council, Ottawa, Canada. May, 2001. "Mobilizing the Semantic Web with DAML-Enabled Web Services."

Workshop on Ontologies in Agent Systems, 5th International Conference on Autonomous Agents, May, 2001. "Mobilizing the Semantic Web with DAML-Enabled Web Services."

Stanford University, Stanford, California. May, 2001. "Mobilizing the Web with DAML-Enabled Web Services."

DARPA DAML-Horus Meeting, Washington, DC. February, 2001. "The Prospects for DAML-Enabled Web Services."

SRI, Menlo Park, California. May, 1999. "Modeling and Diagnosing Dynamical Systems."

Queen's University, Kingston, Ontario. April, 1999. "Modeling and High-Level Control of Physical Systems."

Stanford University, Stanford, California. March, 1999. "Integrating Actions into Diagnostic Problem Solving."

Stanford University, Stanford, California. October, 1998. "Explanatory Diagnosis: Conjecturing Actions to Explain Observations."

NASA Ames Research Center, Mountainview, California. April, 1997. "Representing Actions and State Constraints in Model-Based Diagnosis."

Stanford University, Stanford, California. December, 1996. "SD + Actions: New Representation Problems for Model-Based Diagnosis."

University of Waterloo, Waterloo, Ontario. April, 1996. "Incorporating Actions into Diagnostic Problem Solving."

AT&T Bell Labs, Murray Hill, New Jersey. May, 1995. "Incorporating Actions into Diagnostic Problem Solving."

Xerox Philips Research Center, Webster, New York. May, 1995. "Incorporating Actions into Diagnostic Problem Solving."

Xerox PARC, Palo Alto, California. April, 1995. "Incorporating Actions into Diagnostic Problem Solving."

Simon Fraser University, Burnaby, British Columbia. April, 1995. "Incorporating Actions into Diagnostic Problem Solving."

University of Central Florida, Orlando, Florida. January, 1995. "Towards a Logical Account of Diagnostic Problem Solving."

German Research Centre for Artificial Intelligence (DFKI), Saarbrücken, Germany. May, 1994. "Generating Tests Using Abduction."

Canadian Association of Production and Inventory Control (CAPICS), Calgary, Canada. January, 1988. "Artificial Intelligence in Manufacturing."

INVITED PANELS

Panel member, Panel discussion on Programmatic Representations for Agent Learning, ICML 2025 Workshop on Programmatic Representations for Agent Learning, Vancouver, Canada, July 18, 2025.

Panel member, Panel discussion on Generalization in Planning, AAAI 2025 Workshop on Generalization in Planning, Philadelphia, USA, March 4, 2025.

Panel member, UofT 2024-25 Discovery Series: Conversations with Students about AI, University of Toronto, Toronto, Canada, September 23, 2024.

Panel member, Technophilosophy September Soiree: Can we ensure AI is safe? Isabel Bader Theatre, Toronto, Canada, September 10, 2024.

Panel member, AI Safety Panel at Deep Learning/Reinforcement Learning Summer School, Toronto, Canada, July 9, 2024.

Panel member, Implications of LLMs, the 38th Annual AAAI Conference on Artificial Intelligence (AAAI), Vancouver, Canada, February 24, 2024.

Panel member, Responsible AI for SMBs: Insights from AI leaders, Collision Conference, Toronto, Canada, June 27, 2023.

Panel member, Business Development Bank of Canada (BDC) Responsible AI Panel. AI in Financial Services: "Adopt. Innovate. Regulate?" Toronto, Canada, July 14, 2022.

Panel member, Teaching Ethical and Humanistic Thinking in Computer Science, Applied Ethics Symposium, Canadian Philosophical Association (CPA) Congress 2022, May 18, 2022.

Panel member, Teaching AI Ethics and Responsible AI, EAAI Symposium for Educational Advances in Artificial Intelligence (EAAI), held in conjunction with AAAI, Feb 26, 2022.

Panel member, Responsible AI, 3rd Annual Rotman CPA Ontario Centre for Accounting Innovation Research Conference, Rotman School of Management, University of Toronto, Toronto, Canada, February 17, 2021.

Panel Chair, Advances in RL, AICAN 2021, Canada CIFAR AI Chairs Workshop, January 21, 2021.

Panel member, Ethics in AI Panel, Canadian Undergraduate Conference on AI (CUCAI), Toronto, March 8, 2020.

Panel member, Girls in Tech Conference, Toronto, Canada, April 13, 2019.

Panel member, Vector Institute Research Symposium and Job Fair, Toronto, Canada, February 22, 2019.

Panel member, IJCAI 2018 Doctoral Consortium Career Panel, Stockholm, Sweden, July 13, 2018.

Career Panel member, Increasing Diversity in AI Workshop at the 31st AAAI Conference on Artificial Intelligence (AAAI), San Francisco, USA, February 4, 2017.

Panel member, Global Smartweek Toronto IoT & Machine Intelligence Conference, Toronto, Canada, November 4, 2015.

Panel member, DCS50 Anniversary Launch Panel, University of Toronto, November 13, 2014.

Career Panelist and Mentor, CRA-W/CDC Broadening Participation in AI, 29th Conference on Artificial Intellligence (AAAI 2014), Quebec City, Quebec, July 28, 2014.

Panel member, AI and the Movie Transcendence Panel, sponsored by the University of Toronto, Undergraduate AI Group (UAIG), May 6, 2014.

Panel member, Science at the Movies: Short Circuit, Toronto, ON., March, 2013.

Panel member, Careers in Academia and Industry, Ontario Celebration of Women in Computing (ONCWIC), Kingston, ON., October, 2010.

Expert Panel on Web Services, XML One/Web Services One Conference, San Jose, CA., October, 2001. "The Future of Web Services and Implications for Web Service Developers."

IJCAI-01 Semantic Web Panel, Seventeenth International Joint Conference on Artificial Intelligence (IJCAI-01), Seattle, WA. August, 2001. “The Semantic Web Elephant. What the Blind Men See.”

NRAC-01 Nonmonotonic Reasoning, Action and Change Workshop Panel, Fourth Workshop on Nonmonotonic Reasoning, Action, and Change (NRAC-01), Seattle , WA., August, 2001. “Challenge Problems for Nonmonotonic Reasoning.”

Panel member “The Future of Causation”. AAAI Spring Symposium on Prospects for a Commonsense Theory of Causation, Stanford, California, March, 1998.