Research Statement

Akshat Kumar
School of Information Systems
Singapore Management University
akshatkumar@smu.edu.sg

Background

As our society and urban environments are rapidly becoming connected, they present an opportunity to deploy autonomous agents—from personal digital assistants to self-driving cars—that promise to radically improve productivity and safety, while reducing human efforts and risks. As an example, consider autonomous-taxi fleet optimization where a key problem is to position taxis strategically based on their changing local environment (total demand, other taxis in the same city zone). Similarly, autonomous ships, already on the horizon, are predicted to disrupt the maritime shipping industry. For the world's busiest ports such as Singapore's, a key problem is how to ensure a coordinated movement of ships that enables just-in-time arrival and maintains safety of navigation to avoid environmental disasters such as oil spillage.

With such rapid proliferation of multiagent systems, following challenges and opportunities arise:

- *Scale.* Connected urban environments would require coordinating very large number of agents (e.g., autonomous taxis, vessels, drones). Multiagent coordination approaches need to scale up to such massive systems. However, with scale, there are several opportunities to exploit symmetries such as agent interactions depending only on the aggregate counts of agents, rather than their identities (e.g., congestion, revenue, safety).
- *Decentralized control.* A central control of a large multiagent system creates a bottleneck for future scaling, communication overheads, and is not robust to failures. Therefore, there is a need to increase decentralization of control wherein agents operate mostly autonomously acting based on their local observations of the environments, and only requiring limited communications with a central entity.
- *Resource constrained optimization.* Agents often operate in a resource limited environment. E.g., navigable sea-space in busy megaports such as Singapore's is limited; road network limits vehicle count. Achieving effective coordination despite such resource limitations is the key for practical applicability of multiagent systems.
- *From data to decisions.* Instead of using idealized models of the multiagent system, accurate domain simulators need to be constructed from data, which is often noisy and missing. Fortunately, reasoning about a large agent population is possible using the aggregate data, which is often easier to obtain than tracking each agent individually. This allows for multiagent reinforcement learning approaches to enable agents learn to take better decisions by repeatedly interacting with the simulator.

To address above challenges, *my research aims to develops scalable algorithmic techniques for planning and reinforcement learning in rich, formal models of multiagent coordination such as distributed constraint optimization (DCOP) and decentralized partially observable MDPs (Dec-POMDPs).* Previous planning approaches have suffered from poor scalability or reliance on strong assumptions that limited their deployment. Consequently, my work develops a number of general scalable techniques and frameworks for a large class of multiagent planning problems of practical importance by exploiting the structure and symmetries present in urban-scale multiagent systems. Methodologically, my work uses a synthesis of rigorous techniques from multiple sub-areas of AI, and explores novel connections of planning with probabilistic graphical models, machine learning and mathematical optimization.

Within this broad context, figure 1 shows the evolution of my research over time.

Research Areas

A.1 Multiagent Decision Making Under Uncertainty

My early research during the PhD was towards developing scalable algorithms for distributed constraint optimization (DCOP) and decentralized partially observable MDP (Dec-POMDP) models, which have emerged

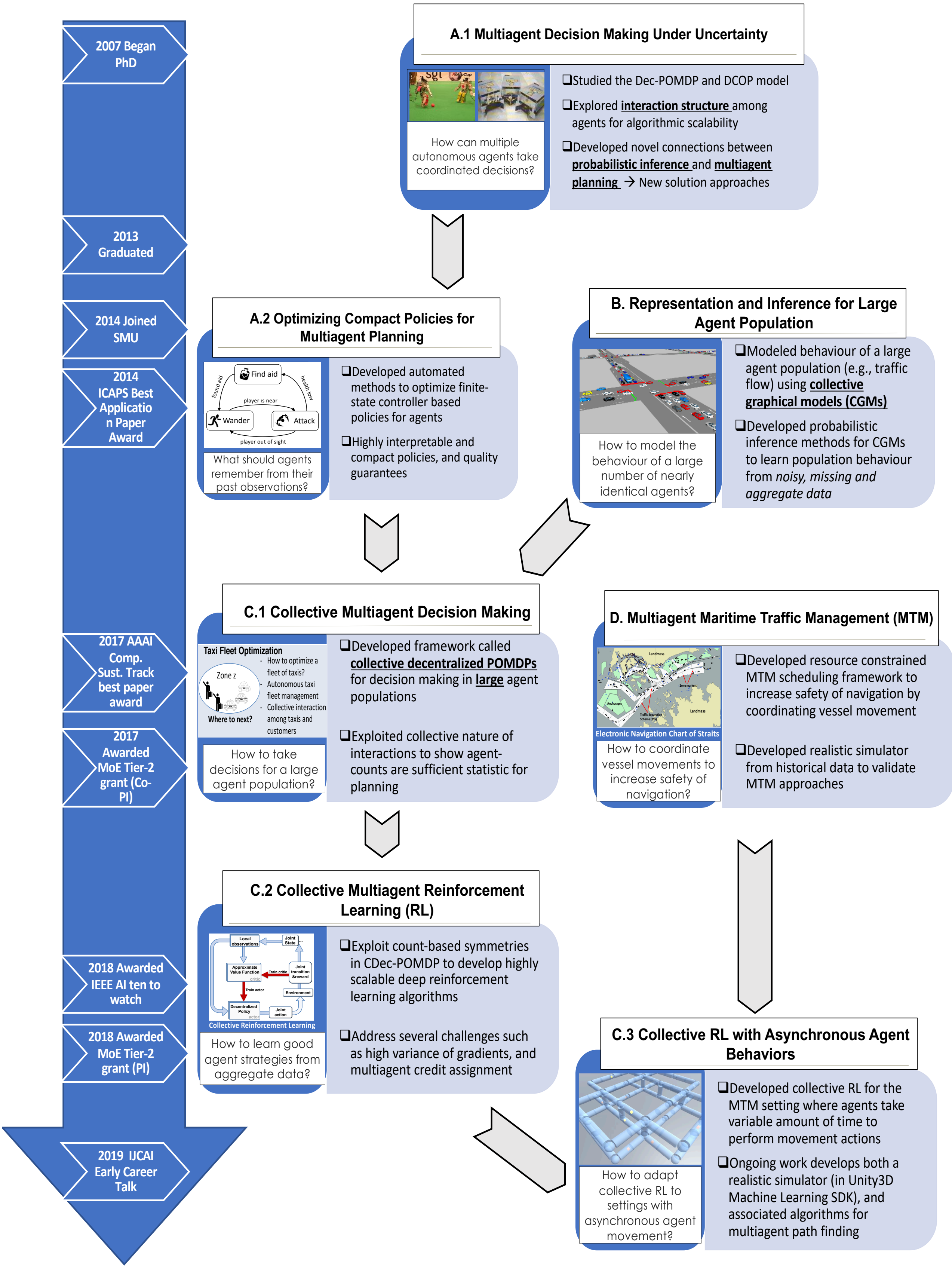

Figure 1: Evolution of my research directions

as popular frameworks for (sequential) cooperative multiagent decision making. These models captures planning problems where agents act based on different partial information about the environment and about each other to maximize a global reward function. I have contributed towards:

- Scalable algorithms for **finite-horizon** Dec-POMDPs where agents operate over a finite plan horizon [7, 8, 11]. I show connections of finite-horizon planning to the problem of constraint optimization (COP) or Markov random fields, a subclass of graphical models. As a result of this connection, established COP solvers can be used to solve Dec-POMDPs resulting in increased algorithmic scalability by multiple orders of magnitude over previous best solvers.
- **Probabilistic inference** based approaches for multiagent planning (for both finite and infinite horizon settings). I have shown how the multiagent planning problem can be reformulated to that of probabilistic inference in a (dynamic) Bayesian network for both the Dec-POMDP [9, 14, 15] and the DCOP models [10]. Using this connection, I show how the problem of likelihood maximization (LM), a popular machine learning problem, is equivalent to that of solving the planning problem optimally. As a result of these connections, a number of approaches from the ML literature become applicable to multiagent planning, leading to a fruitful and productive avenue of research for myself and the wider multiagent planning community. Several state-of-the-art multiagent planning algorithms are based on this planning-as-inference paradigm.
- **Exploiting interaction structure** among agents to scale multiagent planing (increasing scalability from 2 to dozens of agents). This scalability is achieved in settings such as sensor networks where agents are only loosely-coupled. Such restricted agent interactions have been formalized using the network distributed POMDP model [16], and I have developed several state-of-the-art approaches for this model [7, 14, 12].

A.2 Optimizing Compact Policies for Multiagent Planning

Solution approaches for sequential decision making suffer from the *curse of history* which implies agents may need to remember their entire observation history to select the next action. Fortunately, in practice, remembering compact summary of the observation history is sufficient to make good decisions. I have developed mixed-integer programming based methods that optimize both the information that needs be memorized as well as the action to take for each such *memory* using the framework of finite-state controllers [5, 13]. This approach results in *highly compact* policies that are human interpretable, and implementable in resource-limited agents. This work is currently among the state-of-the-art methods for optimizing controller-based policies for 2-agent Dec-POMDPs, and validated on several publicly available benchmarks.

At a high-level, most of my work during the Ph.D. and early tenure at SMU was *model-based* where the knowledge of the complete planning model was assumed. The scalability of algorithms developed was up to **few dozen agents**, which at the time was significant given the NEXP-Hard complexity of a 2-agent Dec-POMDP, and the NP-Hard complexity of DCOPs. During the following years, my focus has shifted towards learning planning models and domain simulators from data, and using such simulators for effective multiagent decision making specially in urban settings with **hundreds to thousands of agents**.

B Representation and Inference for Large Agent Population

After graduation and during my postdoc at IBM Research, I became interested in how to scale multiagent decision making to practical problems by exploiting the structure of agent interactions in a large population. Towards this end, I began working on the framework of *collective graphical models* (CGMs) [22] which allow **learning the behavior of a population of agents** based only on aggregate population-level data **without requiring to track each individual agent**.

Along with my collaborators, I have developed several scalable message-passing approaches for inference in CGMs [20, 24, 6, 21]. Crucially, these inference algorithms allow to **learn the hidden parameters** governing interaction among agents using only **aggregate, noisy and missing data**, which is the norm in real world settings. These algorithms are highly scalable, up to thousands of agents, by exploiting that most of the agents are homogeneous to each other in a large population. We have used CGM-based models to learn crowd movement patterns in a theme park in Singapore as part of the Living Analytics Research Center at SMU, and were able to achieve good accuracy for predicting wait times at popular attractions [3, 4].

I have been part of the *Fujitsu-SMU Urban Computing and Engineering* (UNiCEN) Corp Lab since 2014. Problem settings at UNiCEN are characterized by a need to control a large fleet of agents (most of whom are homogeneous or belong to a small number of *types*). E.g., an autonomous taxi may observe the demand and the count of other taxis in its current zone to decide where to move next to optimize fleet's total revenue; a vessel in Singapore strait may observe the traffic in its current zone, and decide its speed to reach its next target zone to minimize congestion. Such practical contexts have led me to naturally combine CGMs and decision theoretic reasoning to develop new multiagent planning frameworks and algorithms that can model such large multiagent systems as described next.

C.1 Collective Multiagent Decision Making

We have developed a general decision theoretic framework—**collective decentralized partially observable MDP** (CDec-POMDP)—that allows to control the behavior of a population of nearly identical agents operating collaboratively in an *uncertain* and *partially observable* environment [17]. Our key enabling insight is that agent interactions are governed by the aggregate count and types of agents, and do not depend on the specific identities of individual agents, similar to CGMs. This insight makes it possible to construct scalable and general approaches to multiagent modeling, simulation and optimization that are capable of addressing a range of practical problems in urban systems. This addresses shortcomings of previous multiagent planning approaches which are either general but not scalable or scalable but with limited applicability. Using collective interactions, the CDec-POMDP model succinctly captures real world problems, such as autonomous taxi fleet optimization [17] and maritime traffic management [23], involving thousands of agents.

We have also established several basic properties of CDec-POMDPs such as agent count in different states being the sufficient statistic for agent-based simulation and decision making. Such theoretical properties are exploited for scalability by the solution approaches we have developed.

C.2 Collective Multiagent Reinforcement Learning

We have developed a set of principled approaches for collective decision making with a focus on **reinforcement learning (RL) based algorithms** [18, 19]. The novel contribution is the development and validation of learning approaches that **work with count-based data**, which is the key to computational tractability as it allows us to handle large agent populations including thousands of agents. Such a count-based multiagent learning is not possible with standard RL algorithms. Furthermore, we have also addressed the following core challenges for *multiagent* RL: (i) *stability of learning* in the presence of multiple agents, (ii) *scalability* given the combinatorial state and action spaces in collective multiagent decision making problems, and (iii) *multiagent credit assignment* that helps to identify agents that contributed most to the team's success.

We have validated our approaches using **domain simulators based on real historical data** for taxi supply-demand matching, incidence response in Singapore [19], and maritime traffic coordination in Singapore strait [23]. Our approaches are shown to be significantly better (both in scalability and solution quality) than previous best approaches.

D Multiagent Maritime Traffic Management

As part of the UNiCEN, I have worked on projects targeted towards near term operational needs of the vessel traffic control system in collaboration with the Maritime Port Authority (MPA) of Singapore. Using historical location data of all vessels in Singapore strait and its electronic navigation charts, we have built a realistic **maritime traffic simulator** over the last 4 years. We have developed **intelligent scheduling algorithms** that smartly coordinate the vessel traffic in the narrow strait to increase the safety of navigation while minimally affecting the traffic throughput. We have developed constraint programming and operations research based approaches that compute safe passage plan for arriving vessels [1]. Results using historical data show that we can reduce the traffic density by 50% over the current levels by incurring only marginal average delays (about 30 minutes per vessel) [1]. We have extended our scheduling based approaches to incorporate realistic maritime navigational features and constraints through consultations with domain experts to make sure that resulting schedules are implementable in real world [2].

C.3 Collective RL with Asynchronous Agent Behaviors. Adapting RL to the maritime domain requires **incorporating several domain constraints** such as vessels cannot stop mid-water, have to maintain a minimum cruising speed, and the uncertainty of navigation in the strait which is affected by weather conditions. As a result of such constraints, agents (or vessels) operate in an asynchronous manner—vessels may navigate the same zone in different amount of time. This is a significant departure from the standard multiagent RL case where agents' actions take same amount of time to finish. Our recent and ongoing work has addressed such constraints by **learning the navigation behavior of vessel movement using historical data, incorporating it into our maritime traffic simulator, and adapting our collective RL approaches to this setting** [23]. Empirical results on synthetic and real world problems show that our approach can significantly reduce congestion while keeping the traffic throughput high [23].

Recognition and Accomplishments

- I have been invited to give a talk at the **International Joint Conference on Artificial Intelligence (IJCAI) 2019 Early Career Track**. The early career speakers are chosen among young researchers with an outstanding publication record and high visibility following nominations by IJCAI area chairs.
- I am selected as one of the **IEEE Intelligent System's AI ten to watch** in 2018. It is a highly competitive call (open worldwide every two years) for early career AI researchers who received their Ph.D. in or after 2011. My contribution was titled *Planning and Inference for Multiagent Systems.*
- **Best papers.** •Received **Best paper award** at the AAAI Conference on Artificial Intelligence, Computational Sustainability track, 2017. AAAI is the flagship conference for the AI community. •Received **outstanding Application Paper Award** at the International Conference on Automated Planning and Scheduling (ICAPS), 2014. The ICAPS is the flagship conference for the automated planning and scheduling community.
- I was the co-chair of the **Planning and Learning Track** for the 2018 ICAPS and the co-chair of the doctoral consortium at ICAPS 2019.
- **External grants.** I have received multiple external highly competitive grants by the Ministry of Education (MoE) for their tier-2 program—one as Co-PI (Jan'17-Dec'19, $674,046), and the other as PI (Jan'19-Dec'21, $468,700). The latest grant is towards collective multiagent decision making and RL, my recentmost research topic.
- **Dissertation awards:** Received *best Dissertation Award* at ICAPS 2014, *IFAAMAS Victor Lesser Distinguished Dissertation Runner-up Award* at the International Conference on Autonomous Agents and Multiagent Systems (AAMAS), 2013, and the *best dissertation award* at the school of computer science at the University of Massachusetts Amherst.

Future Directions

I am particularly excited by the opportunities and research questions that are going to arise in our rapidly networked and connected urban environments. To benefit from such connectivity, new computational technologies are needed for increasing our productivity, safety and efficiency. I plan to address in near future:

- **Scalability:** Given very large number of agents (in thousands) in our urban and societal settings, how can we develop coordination models/algorithms that scale well to such large planning problems?
- **Cooperation and Competition:** Real world systems are neither fully cooperative nor fully self-interested. How can we develop novel models and algorithms that can encapsulate such cooperative-competitive behavior of agents in a unified decision theoretic model?
- **Multiagent RL and urban system optimization:** My current work explores deep multiagent RL for urban system optimization. Given the success of single-agent RL, developing such techniques for multiple agents presents several challenges such as nonstationary environment due to multiple learning agents, learning with partial observability, and multiagent credit assignment.
- **Humans and AI:** Humans would always cohabitate with AI based technologies such as human driven cars and autonomous taxis sharing the road, autonomous drones making delivery to our houses. A key question to explore is how to design "safe" algorithms that enable productive cohabitation among humans and AI agents possible.

References

[1] L. Agussurja, A. Kumar, and H. C. Lau. Resource-constrained scheduling for maritime traffic management. In *AAAI Conference on Artificial Intelligence*, pages 6086–6093, 2018.

[2] S. Bhatnagar, A. Kumar, and H. C. Lau. Decision making for improving maritime traffic safety using constraint programming. In *International Joint Conference on Artificial Intelligence (IJCAI)*, 2019.

[3] J. Du, A. Kumar, and P. Varakantham. On understanding diffusion dynamics of patrons at a theme park. In *International conference on Autonomous Agents and Multi-Agent Systems*, pages 1501–1502, 2014.

[4] J. Du, P. Varakantham, A. Kumar, and S. Cheng. Learning and controlling network diffusion in dependent cascade models. In *IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology*, pages 336–343, 2015.

[5] A. Kumar, H. Mostafa, and S. Zilberstein. Dual formulations for optimizing Dec-POMDP controllers. In *International Conference on Automated Planning and Scheduling*, pages 202–210, 2016.

[6] A. Kumar, D. Sheldon, and B. Srivastava. Collective diffusion over networks: Models and inference. In *Conference on Uncertainty in Artificial Intelligence, UAI*, 2013.

[7] A. Kumar and S. Zilberstein. Constraint-based dynamic programming for decentralized POMDPs with structured interactions. In *International Conf. on Autonomous Agents and Multiagent Systems (AAMAS)*, pages 561–568, 2009.

[8] A. Kumar and S. Zilberstein. Event-detecting multi-agent MDPs: Complexity and constant-factor approximation. In *IJCAI*, pages 201–207, 2009.

[9] A. Kumar and S. Zilberstein. Anytime planning for decentralized POMDPs using expectation maximization. In *International Conf. on Uncertainty in Artificial Intelligence (UAI)*, pages 294–301, 2010.

[10] A. Kumar and S. Zilberstein. MAP estimation for graphical models by likelihood maximization. In *In Neural Information Processing Systems*, pages 1180–1188, 2010.

[11] A. Kumar and S. Zilberstein. Point-based backup for decentralized POMPDs: Complexity and new algorithms. In *International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, pages 1315–1322, 2010.

[12] A. Kumar and S. Zilberstein. Message-passing algorithms for large structured decentralized POMDPs. In *International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, pages 1087–1088, 2011.

[13] A. Kumar and S. Zilberstein. History-based controller design and optimization for partially observable MDPs. In *International Conference on Automated Planning and Scheduling*, pages 156–164, 2015.

[14] A. Kumar, S. Zilberstein, and M. Toussaint. Scalable multiagent planning using probabilistic inference. In *International Joint Conf. on Artificial Intelligence (IJCAI)*, pages 2140–2146, 2011.

[15] A. Kumar, S. Zilberstein, and M. Toussaint. Probabilistic inference techniques for scalable multiagent decision making. *Journal of Artificial Intelligence Research*, 53:223–270, 2015.

[16] R. Nair, P. Varakantham, M. Tambe, and M. Yokoo. Networked distributed POMDPs: A synthesis of distributed constraint optimization and POMDPs. In *Proc. of the Twentieth National Conference on Artificial Intelligence*, pages 133–139, 2005.

[17] D. T. Nguyen, A. Kumar, and H. C. Lau. Collective multiagent sequential decision making under uncertainty. In *AAAI Conference on Artificial Intelligence*, pages 3036–3043, 2017.

[18] D. T. Nguyen, A. Kumar, and H. C. Lau. Policy gradient with value function approximation for collective multiagent planning. In *Advances in Neural Information Processing Systems*, pages 4319–4329, 2017.

[19] D. T. Nguyen, A. Kumar, and H. C. Lau. Credit assignment for collective multiagent RL with global rewards. In *Advances in Neural Information Processing Systems*, pages 8113–8124, 2018.

[20] D. T. Nguyen, A. Kumar, H. C. Lau, and D. Sheldon. Approximate inference using DC programming for collective graphical models. In *International Conference on Artificial Intelligence and Statistics, AISTATS*, pages 685–693, 2016.

[21] D. Sheldon, T. Sun, A. Kumar, and T. G. Dietterich. Approximate inference in collective graphical models. In *International Conference on Machine Learning*, pages 1004–1012, 2013.

[22] D. R. Sheldon and T. G. Dietterich. Collective graphical models. In *Advances in Neural Information Processing Systems*, pages 1161–1169, 2011.

[23] A. J. Singh, D. T. Nguyen, A. Kumar, and H. C. Lau. Multiagent decision making for maritime traffic management. In *AAAI Conference on Artificial Intelligence*, pages 6171–6178, 2019.

[24] T. Sun, D. Sheldon, and A. Kumar. Message passing for collective graphical models. In *International Conference on Machine Learning, ICML*, pages 853–861, 2015.