Human Factors in Robots, Drones and Unmanned Systems, Vol. 93, 2023, 183–188
https://doi.org/10.54941/ahfe1003761

Multiple Agents Interacting via Probability Flows on Factor Graphs

Francesco A. N. Palmieri[1,2], Krishna R. Pattipati[2], Giovanni Di Gennaro[1], Amedeo Buonanno[3], and Martina Merola[1]

[1]Dipartimento di Ingegneria, Università degli Studi della Campania Luigi Vanvitelli, via Roma 29, 81031 Aversa (CE), Italy
[2]Department of Electrical and Computer Engineering, University of Connecticut, Storrs, CT 06269, USA
[3]Department of Energy Technologies and Renewable Energy Sources, Italian National Agency for New Technologies, Energy and Sustainable Economic Development (ENEA), Portici, 80055 Naples, Italy

ABSTRACT

Factor graphs with forward-backward probability propagation flows are used to model interacting agents. Each agent follows an MDP (Markov Decision Process) on a graph with the probability flow that is partially shared with the others. Each MDP bases its decisions on current knowledge and future predictions about itself and of the others, generating a complex time-varying scenario. The paper reports some preliminary results on the use of the sum-product algorithm applied to the interacting multiple agent model, where agents with individual destination goals move on a small rectangular grid with obstacles that need to be avoided, thereby necessitating coordination among the agents. Non-trivial solutions and behaviors are observed in the presence of conflicting paths and objectives.

Keywords: Factor graphs, Probability propagation, Path planning, Multiple agents, Interacting Markov decision processes

INTRODUCTION

Studies in expert team decision-making demonstrate that effective teams have shared goals and shared mental models to coordinate implicitly and/or explicitly with minimal communication. They can establish trust through cross-training, and are matched to the task structure through planning (Kleinman et al., 1990; Cannon-Bowers et al., 1993; Blickensderfer et al., 2010; Rico et al., 2008). The key questions then are the following: Do the best practices of human teams translate to hybrid teams comprised of human and AI agents or autonomous agents alone? Is there a mathematical framework for systematically studying shared goals and shared mental models?

This paper proposes a factor graph-based mathematical framework for studying multi-agent interaction in agile cooperative planning situations. We find that one of the most promising avenues for modeling complex multistage decision-making scenarios is to use stochastic approaches, where uncertainties in the dynamic evolution of the environment and imperfect sensor

© 2023. Published by AHFE Open Access. All rights reserved.

observations are described using probabilistic models. Stochastic dynamic programming provides a systematic methodology for modeling and solving multiple interacting Markov Decision Processes (MDPs), wherein each agent (MDP) has partial information about the other agents (MDPs) in the team (Bertsekas, 2021). Each agent then acts by accounting for both its own objectives and the anticipated behaviors of others, as if they have a shared mental model (Palmieri et al., 2022; Di Gennaro et al. 2022). Our model could represent agent roles in a team and interdependencies of agent behaviors in accomplishing team goals. Lack of consideration of team-level interdependencies and overreliance on automation can result in fatal collisions, as documented in National Transportation Safety Board reports (NTSB, 2022).

THE BAYESIAN MODEL

Figure 1 shows the conceptual model in which agents keep in their minds both their own planning model and their best predictions of the others.

We have demonstrated in (Palmieri et al., 2022) that, using probabilistic factor graphs (Palmieri, 2016; Di Gennaro et al., 2019), we can solve the single-agent MDP stochastic control problems using message propagation rules that subsume Dynamic Programming, Maximum likelihood, Maximum entropy, and Free-energy-based methods. In this contribution, we show how this framework can be extended to model multiple agents using interacting factor graphs that exchange probability distributions among themselves. Bertsekas (2021) has already considered Dynamic Programming for multiple interacting systems, but our framework is more general in that it can be adapted to various cost functions by simply changing some of the propagation rules (Palmieri et al. 2022). Figure 2 shows a section of our MDP factor graph for a generic agent i at time t. Details about the blocks defined in the picture can be found in (Palmieri et al., 2022), where the system evolution is governed by conditional transition probabilities (bottom blocks). The prior blocks that account for rewards, when translated into probabilities, inject messages into the MDP graph to model constraints, such as obstacles, predefined semantic areas (e.g., grassy areas, pathways) and goals. The evolution of message propagation determines the decision an agent takes at time step t, based on accumulated information coming from the future via backward messages, propagated back in time from the terminal time T (T can be large).

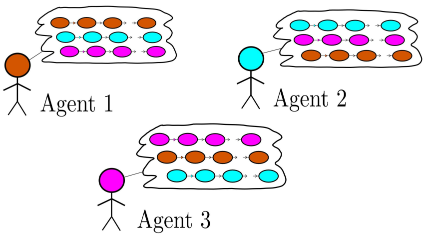

Figure 1: Each agent keeps a running set of interacting models for itself and for the others that extends from the present to a foreseeable future.

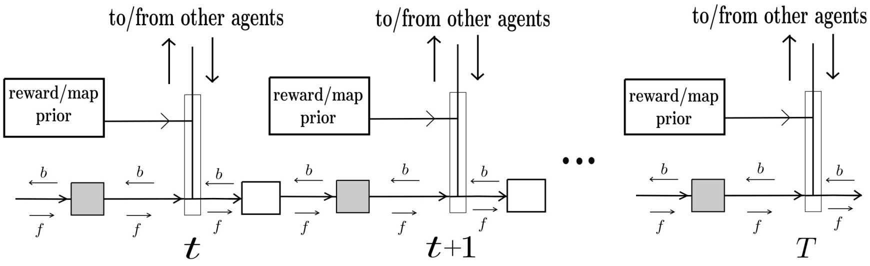

Figure 2: A section of the factor graph for agent *i*. The bottom blocks are the conditional transition probabilities. For more details, please refer to (Palmieri et al., 2022).

In a multi-agent scenario, there is a similar model for each agent with possibly a different destination (goal) and a different reward structure. In addition, the probability flow, unrolled in time, is shared with other agents. Recall that, in belief networks, an outgoing message distribution is the "summary" of the information accumulated from the backward and the forward flows. In other words, the outgoing messages are the posterior probabilities of that agent to be in a certain position at a given future time. These are complemented to one and injected into the other agents' message flows. In this way, each agent has a variable-constraint map in its flow that impacts its behavior and that prevents collisions with the other agents. More importantly, since this information is injected into the whole graph unrolled in time, it is expected that predictions are used by each agent for making current decisions that favorably impact the future. Indeed, the interdependence of team member behavior to achieve team goals implies that "the performance of a team is not decomposable to, or an aggregation of, individual performances" (NAP, 2022).

SIMULATIONS

We have begun experimenting with this framework with a limited number of interacting agents (three) on a small rectangular discrete grid with specified starting points and destination goals for each agent, obstacles placed in various positions, with narrow passages, small mazes, etc. The scheduling of the agents is fixed a priori or may change over time, and the forward-backward flow for each agent's MDP is computed at every time step. In our first set of experiments, we have assumed that all agents have complete knowledge about goals and constraints of the others for the remaining planning horizon $[t, T]$. Limitations in time, or imperfect knowledge of agents or of the environment can be easily included in the model and are being explored in our current work.

Figure 3 shows the results of an experiment with three agents on a small 8x8 grid with obstacles. Each agent can move to one of eight neighboring positions or stay still. Each agent has its own goal (circles) and starts from a different initial position.

Figure 3 shows the sequence of steps the agents take at different time frames. In this experiment, the time horizon T was set to 15 and the agents are scheduled in the sequence blue-violet-brown. Note that the brown agent starts heading towards its goal, but, to avoid a potential conflict with the blue agent in going through the same narrow passage way, decides to back up and wait for the blue agent to pass by (Figure 3, t=4 and t=5), before heading again for its goal (Figure 3, from t=6). The violet agent is so far from the others that it never needs to interact with other agents in reaching its goal in just one step. We would like to emphasize that the observed context-dependent behavior is completely a consequence of the probability flow in the system. No hand adjustments have been made.

Figure 4 shows the results of another experiment on a larger 16x16 grid that includes a very narrow corridor where the agents may collide. The time horizon T is 30 here and the agents are scheduled in the same sequence as before: blue-violet-brown. Note how the violet and brown agents are heading at the beginning towards their goals using the same narrow passage (Figure 4, t=5 and t=6). Apparently, their predictions from the future have not been able to tell them otherwise. However, as soon as the probability flow detects a potential conflict two steps away, the brown agent backs up all the way to the end (Figure 4, t=13), to let the violet agent pass by (Figure 4, t=11 and t=13) and then heads to its goal (Figure 4, t=24 and t=28). The blue agent has no interactions with the others in reaching its goal in just one step.

We have been experimenting with other scenarios and, even though our results are preliminary, it appears that the solutions found by the probabilistic model are quite striking. Complex strategies emerge naturally as a result of the probability distributions that flow through the interacting flow graphs. In (Di Gennaro et al., 2022), we have already presented a scheme in which the agents follow their complete paths in a hierarchical order, where a unique value function is pre-computed for every agent. In this contribution, even if there is a scheduling order, the posterior probabilities flow freely in the unrolled time model and are re-computed at every time step.

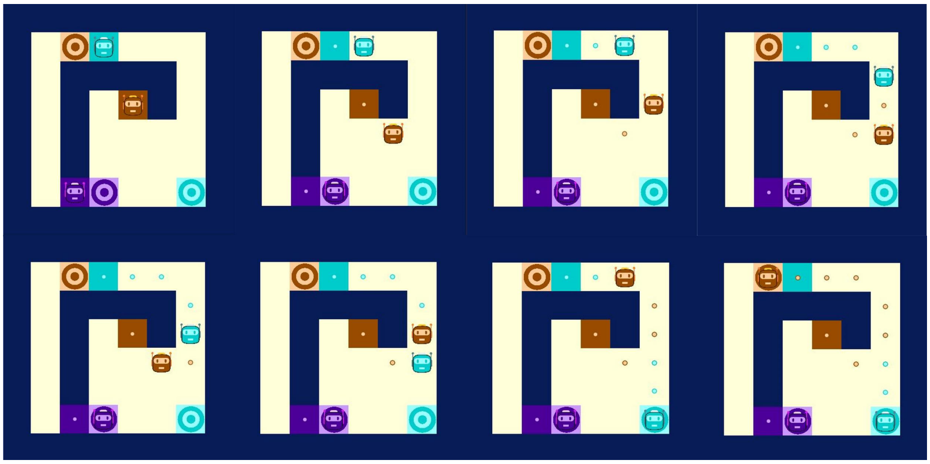

Figure 3: Simulation of three interacting agents on an 8x8 map. The pictures, to be read in lexicographic order, correspond to t = 0, 1, 2, 3, 4, 5, 7, 10 (start at upper left to right in the first row and then from left to right in the second row).

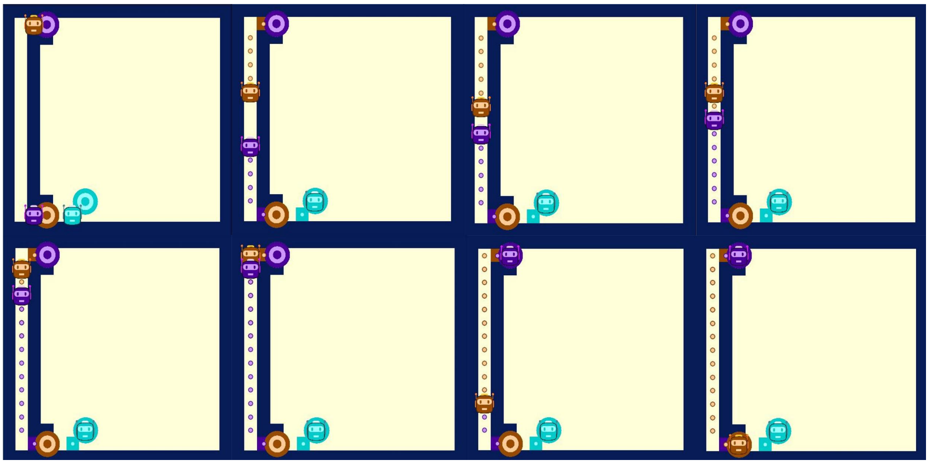

Figure 4: Simulation of three interacting agents on a 16x16 map. The pictures, to be read in lexicographic order, correspond to t = 0, 5, 6, 7, 11, 13, 24, 28.

CONCLUSION

In this short paper, we have presented preliminary results of a very promising framework for handling multiple interacting agents with conflicting goals and obstacles. The probability flow allows great flexibility in tuning the information that each agent has on the others: it can span the whole range that goes from complete knowledge of goals and positions about the others, to a limited probabilistic awareness, both in precision and in time, of where the other agents may be located at future time steps. We are pursuing further work on this framework as it seems to allow systematic addressing of questions, such as minimal amount of information needed for effective team coordination in the face of changes in goals, communication bandwidth, grid parameters and agent status. We also plan to address the team fit (congruence) hypothesis that maximal team performance, measured in terms of speed and precision, accrues when team structures are congruent with the task structure (Levchuk, 2004).

REFERENCES

Bertsekas D. (2021) "Multi-agent Reinforcement Learning: Rollout and Policy Iteration." *IEEE/CAA Journal of Automatica Sinica*, Vol. 8, No. 2. pp. 249–272.

Blickensderfer, E. L., Reynolds, R., Salas, E., & Cannon-Bowers, J. A. (2010) "Shared expectations and implicit coordination in tennis doubles teams." *Journal of Applied Sport Psychology*, Vol. 22, pp. 486–499.

Cannon-Bowers, J. A., Salas, E., & Converse, S. A. (1993). "Shared mental models in expert team decision making." in N. J. Castellan, Jr. (Ed.), *Individual and group decision-making: Current issues*, pp. 221–246, Hillsdale, NJ: Erlbaum.

Di Gennaro, G., Buonanno, A. and Palmieri, F. A. N. (2019) "Optimized realization of Bayesian networks in reduced normal form using latent variable model," *Soft Computing*, Berlin, Germany: Springer, Vol. 25, pp. 7029–7040.

Di Gennaro, G., Buonanno, A., Fioretti, G., Verolla, F., Pattipati, K. R., and Palmieri, F. A. N (2022). "Probabilistic Inference and Dynamic Programming: A Unified

Approach to Multi-Agent Autonomous Coordination in Complex and Uncertain Environments." *Frontiers in Physics*, Vol. 10, pp. 1–12.

Kleinman, D. L., Luh, P. B., Pattipati, K. R. and Serfaty, D. (1990). "Mathematical Models of Team Distributed Decision-making," in R. W. Sweezy and E. Salas (eds.) *Teams: Their Training and Performance*, New York: ABLEX, pp. 177–218.

Levchuk, G. M., Meirina, C., Pattipati, K. R., and Kleinman, D. L. (2004). "Normative Design of Project-based Organizations: Part III - Modeling Congruent, Robust and Adaptive Organizations," IEEE Transactions on SMC-: Part A: Systems and Humans, Vol. 34, No. 3, 2004, pp. 337–350.

NAP (2022). https://nap.nationalacademies.org/read/26355/chapter/4, page 12.

NTSB (2022). https://www.ntsb.gov/Advocacy/safety-topics/Pages/automated-vehicles-investigations.aspx

Palmieri, F. A. N. (2016) "A comparison of algorithms for learning hidden variables in Bayesian factor graphs in reduced normal form," *IEEE Transaction on Neural Networks and Learning Systems*, Vol. 27, no. 11, pp. 2242–2255.

Palmieri, F. A. N., Pattipati, K. R., Di Gennaro, G., Fioretti, G., Verolla, F. and Buonanno, A. (2022). "A Unifying View of Estimation and Control Using Belief Propagation with Application to Path Planning." *IEEE Access*, Vol. 10, pp. 15193–15216.

Rico, R., Sánchez-Manzanares, M., Gil, F., & Gibson, C. (2008). "Team implicit coordination processes: A team knowledge–based approach." *Academy of Management Review*, Vol. 33, pp. 163–184.