



Perkembangan Perilaku Pembelajaran Peserta Didik dengan Menggunakan Partially Observable Markov Decision Processes

Maxtulus Junedy Nababan

Universitas Quality, Indonesia

Email : maxtnab@gmail.com

Abstract *The School Community can be considered as an agent in a multi-agent system because there is dynamic interaction in a school system. The important thing in a multi-agent system is managing relationships between agents to achieve coordinated behavior by developing knowledge, attitudes and practices, namely the factors that shape environmental behavior. This research extends the framework of partially observable Markov decision processes (POMDPs) to multi-agent settings by incorporating the idea of agent models into the state space. The results of this research show that (POMDP) is effective for understanding and managing learning behavior in schools.*

Keyword : Markov Model, Learning Behaviours, Multi-Agent

Abstrak Komunitas sekolah dapat dianggap sebagai agen dalam sistem multiagen karena terdapat interaksi yang dinamis dalam sistem sekolah. Hal penting dalam sistem multiagen adalah mengelola hubungan antar agen untuk mencapai perilaku yang terkoordinasi dengan mengembangkan pengetahuan, sikap, dan praktik, yaitu faktor-faktor yang membentuk perilaku lingkungan. Penelitian ini memperluas kerangka kerja proses keputusan Markov yang dapat diamati sebagian (POMDP) ke pengaturan multiagen dengan menggabungkan gagasan model agen ke dalam ruang keadaan. Hasil penelitian ini menunjukkan bahwa (POMDP) efektif untuk memahami dan mengelola perilaku belajar di sekolah.

Kata kunci: Model Markov, Perilaku Belajar, Multiagen

1. LATAR BELAKANG

Komunitas Sekolah dapat dianggap sebagai agen dalam sebuah multi agen sistem karena terdapat interaksi yang dinamis dalam sebuah sistem sekolah. penelitian ini memperluas kerangka proses keputusan Markov yang dapat diamati sebagian (POMDPs) ke pengaturan multi-agen dengan memasukkan gagasan model agen ke dalam ruang keadaan. Lebih khusus lagi, POMDP akan dibangun untuk menggambarkan bagaimana interaksi antara pemangku jabatan dan peserta didik dapat dianalisis untuk memahami proses tersembunyi dari pertumbuhan kompetensi peserta didik dalam konteks pembelajaran.

Dalam tulisan ini diasumsikan bahwa ada agen yang mampu mendukung kinerja unit kerja menuju pencapaian misi. Untuk tujuan itu, dapat dikatakan bahwa terdapat agen otonom yang merupakan entitas cerdas yang bertindak secara rasional dan sengaja sehubungan dengan misi dan status pengetahuannya saat ini (Wooldridge, 1999). Sistem multi-agen (MAS) berkaitan dengan studi tentang agen otonom yang berinteraksi dalam lingkungan yang sama, berbagi sumber daya, bernegosiasi dan berkolaborasi untuk mencapai tujuan mereka, membentuk koalisi, memiliki konflik, dll.

Koordinasi yang efektif sangat penting jika agen otonom ingin mencapai misi dalam sistem multi-agen (MAS). Koordinasi tersebut diperlukan untuk mengelola berbagai bentuk ketergantungan yang secara alami terjadi ketika agen memiliki tujuan yang saling terkait, ketika mereka berbagi lingkungan yang sama, atau ketika ada sumber daya bersama. Koordinasi adalah proses di mana agen beralasan tentang tindakan lokalnya dan tindakan yang diantisipasi orang lain untuk mencoba dan memastikan komunitas bertindak dengan cara yang koheren (Jennings, 1996). Isu yang sangat penting dalam merancang teknik koordinasi untuk sistem multi-agen adalah mengelola situasi di mana tindakan satu agen mempengaruhi operasi agen lain tindakan (Malone & Crowston, 1994). Beberapa upaya penggunaan representasi eksplisit CR telah menunjukkan kegunaannya untuk mencapai koordinasi dalam MAS. Namun, studi ini telah mempertimbangkan lingkungan yang diketahui sepenuhnya, hubungan statis/stabil (Lesser et al., 2004) dan tidak ada/sedikit pembelajaran (D. Jensen et al., 1999).

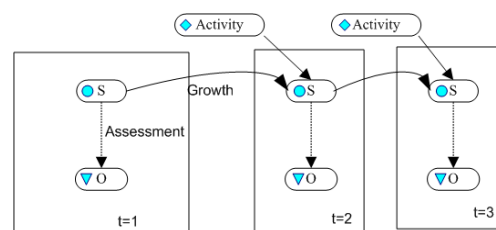
Proses keputusan Markov sebagian yang dapat diamati secara interaktif partially observable Markov decision processes (I-POMDP) menyediakan kerangka kerja untuk pengambilan keputusan berurutan di lingkungan multiagen yang dapat diamati sebagian. Mereka menggeneralisasi POMDP ke pengaturan multiagen dengan memasukkan model yang dapat dihitung agen lain di ruang keadaan bersama dengan keadaan lingkungan fisik. Model mencakup semua informasi yang mempengaruhi perilaku agen, termasuk preferensi, kemampuan, dan keyakinan mereka, dan dengan demikian analog dengan jenis dalam permainan Bayesian. I-POMDP mengadopsi pendekatan subjektif untuk memahami perilaku strategis, yang berakar pada kerangka teori keputusan yang mengambil perspektif pembuat keputusan dalam interaksi.

2. KAJIAN TEORITIS

Asumsi Markov ini memungkinkan penilaian periodik dan proses keputusan instruksional dimodelkan sebagai proses keputusan Markov (MDP) (T. Li et al., 2021). Asumsi Markov mungkin tidak lama jika ada beberapa latensi untuk proses memperoleh kemahiran, atau jika ada semacam efek momentum (misalnya, baru saja mempelajari suatu konsep membuat siswa lebih reseptif untuk mempelajari konsep berikutnya). Sering kali suatu proses dapat diubah menjadi proses Markov dengan menambahkan variabel lain ke status saat ini untuk menangkap ketergantungan pada masa lalu.

Gambar 1 menunjukkan beberapa irisan waktu MDP untuk pembelajaran siswa. Setiap irisan waktu, $t = 1, 2, 3, \dots$, mewakili periode waktu yang singkat di mana kemahiran pelajar dianggap konstan. Variabel S_t (vektor bernilai untuk mencerminkan banyak keterampilan)

mewakili kemahiran siswa pada setiap irisan waktu. Variabel O_t (vektor dinilai untuk mewakili beberapa tugas atau beberapa aspek di mana hasil tugas dapat dinilai) mewakili hasil yang diamati dari tugas penilaian yang diberikan selama irisan waktu itu. Kegiatan yang dipilih oleh instruktur dan terjadi antara irisan waktu. Untuk memodelkan penilaian formatif, irisan waktu dapat mewakili peristiwa periodik di mana penilaian *benchmark* diberikan (misalnya, kuartal tahun akademik). Untuk memodelkan sistem pembelajaran elektronik, irisan waktu dapat mewakili periode di mana sistem berada dalam mode penilaian, sedangkan aktivitas mewakili tugas instruksional. Sebagai pendekatan penyederhanaan, asumsikan bahwa tidak ada pertumbuhan —perubahan status variabel kecakapan—terjadi dalam irisan waktu (Metelli et al., 2018).



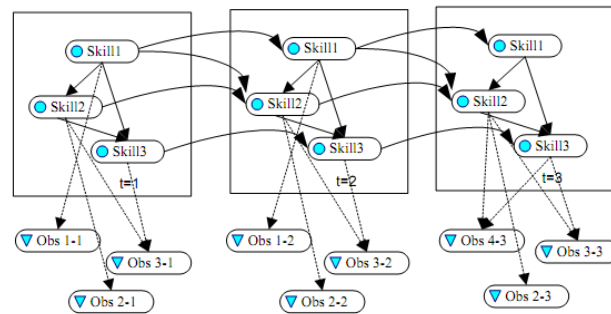
Gambar 1: Pengujian sebagai proses keputusan Markov

Model ini menunjukkan MDP yang terdiri dari dua submodel (akan ditentukan kemudian), satu untuk penilaian dan satu untuk pertumbuhan. Node yang ditandai dengan lingkaran (O) mewakili variabel laten, node yang ditandai dengan segitiga (∇) adalah kuantitas yang dapat diamati, dan node yang ditandai dengan berlian ($\diamond$) adalah variabel keputusan yang kuantitasnya ditentukan oleh instruktur.

Secara umum, variabel profisiensi S tidak dapat diamati, oleh karena itu Gambar 1 mewakili proses keputusan Markov yang dapat diamati sebagian (POMDP) (Boutilier et al., 1999). Karena penyelesaian POMDP umumnya sulit, peluang harus dicari untuk memanfaatkan struktur khusus apa pun dalam model yang dipilih untuk penilaian dan pertumbuhan. Secara khusus, hubungan antara kemahiran dapat dimodelkan dengan jaringan Bayesian.

Jaringan Bayesian Dinamis (DBNs) adalah jaringan Bayesian di mana beberapa tepi mewakili perubahan dari waktu ke waktu (Dean & Kanazawa, 1989). (Pada Gambar 3, tepi melengkung bersifat temporal sedangkan tepi lurus menghubungkan node dalam satu irisan waktu.) DBN pada dasarnya adalah jaringan Bayesian yang diulang pada setiap irisan waktu dengan tepi tambahan yang menghubungkan (kemahiran) variabel pada satu titik waktu ke variabel di titik waktu berikutnya. (Boutilier et al., 1999) mencatat bahwa dapat dibangun dengan terlebih dahulu membangun jaringan Bayesian dasar untuk memodelkan titik waktu

awal dan kemudian membangun jaringan Bayesian dua irisan waktu (2TBN) untuk memodelkan perubahan dari waktu ke waktu.



Gambar 2: Jaringan Bayesian dinamis untuk Pendidikan

Dalam gambar ini, status variabel profisiensi pada setiap irisan waktu direpresentasikan dengan jaringan Bayesian. Tepi putus-putus mewakili pengamatan potensial (melalui tugas penilaian) pada setiap irisan waktu. Tepi melengkung mewakili hubungan temporal yang memodelkan pertumbuhan di antara keterampilan. Perhatikan bahwa tugas yang diberikan dan yang dapat diamati dapat berubah dari titik waktu ke titik waktu.

Meskipun DBN telah dikenal selama lebih dari satu dekade, sampai saat ini kompleksitas komputasinya membuat mereka tidak praktis untuk digunakan dalam konteks otentik. (Boyen & Koller, 2013) mempresentasikan teknik aproksimasi untuk menarik kesimpulan dari DBN. Baik (Koller & Lerner, 2001) dan (Murphy & Russell, 2001) menyajikan perkiraan berdasarkan penyaringan partikel (Doucet et al., 2001; Liu, 2001). (Takikawa et al., 2012) menyarankan efisiensi lebih lanjut dalam algoritma dengan mencatat variabel yang tetap dari waktu ke waktu, menghasilkan jaringan Bayesian yang sebagian dinamis.

Ada korespondensi yang erat antara model POMDP Gambar 1 dan objek desain yang berpusat pada bukti (Mislevy et al., 2003). Node S_t mewakili model profisiensi dan node O_t (atau lebih tepatnya tepi antara S_t dan O_t) mewakili model bukti. Apa yang ditambahkan oleh kerangka kerja ini adalah model tindakan: 2TBN yang mewakili perubahan antara S_{t1} dan S_{t2} . Tantangan dari penelitian ini adalah untuk membangun model matematika sederhana dari pertumbuhan kecakapan (2TBN) yang sesuai dengan pemahaman kognitif dari proses. Istilah model tindakan mengakui bahwa tindakan yang diambil oleh instruktur akan mempengaruhi tingkat kecakapan masa depan siswa.

3. METODE PENELITIAN

Model *single agent* (agen tunggal) dapat dihitung dengan menurunkan POMDP yang didefinisikan sebagai berikut.

$$POMDP_i = \langle S, A_i, T_i, \Omega_i, O_i, R_i \rangle$$

- Dimana S = rangkaian kondisi lingkungan yang ada
- A_i = rangkaian tindakan agen i yang dapat dilakukan.
- T_i = fungsi transisi $T_i : S \times A_i \times S \rightarrow [0, 1]$ yang menggambarkan hasil dari tindakan agen i .
- Ω_i = rangkaian observasi yang dilakukan agen i .
- O_i = fungsi observasi agen $O_i : S \times A_i \times \Omega_i \rightarrow [0, 1]$ yang merupakan probabilitas observasi jika agen melakukan bermacam tindakan yang menyebabkan perubahan kondisi atau kondisi yang berbeda dari sebelumnya.
- R_i = fungsi reward yang mewakili karakteristik agent i . ($R_i : S \times A_i \rightarrow \mathbf{R}$).

Penyelesaian yang diharapkan dari POMDP ini adalah petunjuk ataupun prosedur yang memetakan observasi sejarah ataupun langkah-langkah perilaku menuju tindakan. Apabila POMDP digunakan untuk *single agent* dengan demikian maka bagi komunitas sekolah diperlukan sistem *multi-agent* dengan menurunkan I-POMDPs (*Interactive POMDP*). Pendekatan ini menyediakan sebuah kerangka urutan pengambilan keputusan di dalam lingkungan *multi-agent*. Model ini mencakup semua informasi mempengaruhi perilaku agen yang bersangkutan termasuk preferensi mereka, kemampuan dan keyakinan mereka. Dalam penelitian ini yang dimaksud dengan preferensi adalah yang diasumsikan sebagai dorongan internal dan dorongan eksternal dan motivasi. Sementara itu kemampuan (*capabilities*) diasumsikan sebagai pengetahuan, dan keyakinan diasumsikan sebagai persepsi, sikap dan ekspektasi.

Ketika membahas mengenai keadaan *multi-agent*, ini artinya proses generalisasi I-POMDPs, tetapi untuk meyakinkan agar model ini dapat dihitung maka diasumsikan sekumpulan tabulasinya adalah terbatas sehingga agen dapat memaksimalkan ekspektasi pencapaiannya dalam kurun waktu tertentu. Generalisasi dari POMDP adalah I-POMDP dari *multi-agent* i , yaitu $I-POMDP_{i,l} = (S_{i,l}, T_i, \Omega_i, O_i, R_i)$

4. HASIL

Variabel-variabel dalam penelitian ini adalah sebagai berikut:

- Pengetahuan

- Sikap
- Praktik
- Peran Kepala Sekolah ($S_{\Omega A}$)
- Peran Guru ($S_{\Omega A}$)
- Peran Orang Tua ($S_{\Omega A}$)
- Peran Komite Sekolah ($S_{\Omega A}$)

Dalam penelitian ini, pendekatan POMDP digunakan untuk mendapatkan nilai probabilitas untuk transisi KAP siswa (Boutilier et al., 1999; Hauskrecht, 2000; Kaelbling et al., 1998; Monahan, 1982).

- Dalam keadaan i , POMDP didefinisikan sebagai $POMDP_i = \langle S_i, A_i, T_i, \Omega_i, O_i, R_i \rangle$
- Di mana S_i adalah urutan kondisi lingkungan yang ada.
- A_i adalah urutan tindakan yang dapat dilakukan oleh agen i .
- T_i adalah fungsi transisi $-T_i: S \times A_i \times S \rightarrow [0,1]$ yang menggambarkan hasil tindakan agen i .
- Ω_i adalah urutan pengamatan yang dilakukan oleh agen i .
- O_i adalah fungsi pengamatan dari agen $-O_i: S \times A_i \times \Omega_i \rightarrow [0,1]$ yang merupakan probabilitas pengamatan jika agen mengambil berbagai tindakan yang menyebabkan perubahan kondisi atau kondisi yang berbeda dari sebelumnya.
- R_i adalah fungsi penghargaan yang mewakili karakteristik agen i ($R_i: S \times A_i \rightarrow R$).

Peran agen dan KAP (*Knowledge, Attitude, Practice*) siswa beserta komunitas sekolah di dapat ditemukan sebagaimana tercantum dalam tabel berikut.

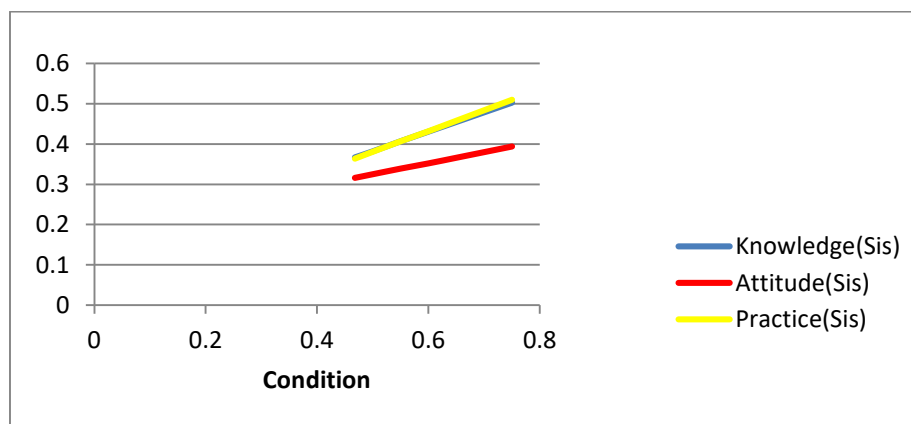
Tabel 1 . Peran Agen dan KAP Siswa

No	Deskripsi	Persentase
1	Peran Kepala Sekolah	36,77
2	Peran Guru	45,12
3	Peran Komite Sekolah	35,15
4	Peran Orang Tua	45,08
5	Pengetahuan Siswa (<i>Knowledge</i>)	26,96
6	Sikap Siswa (<i>Attitude</i>)	56,49
7	Praktik Siswa (<i>Practice</i>)	66,94

Tabel 2. Distribusi data berdasarkan variabel S , Ω , dan A

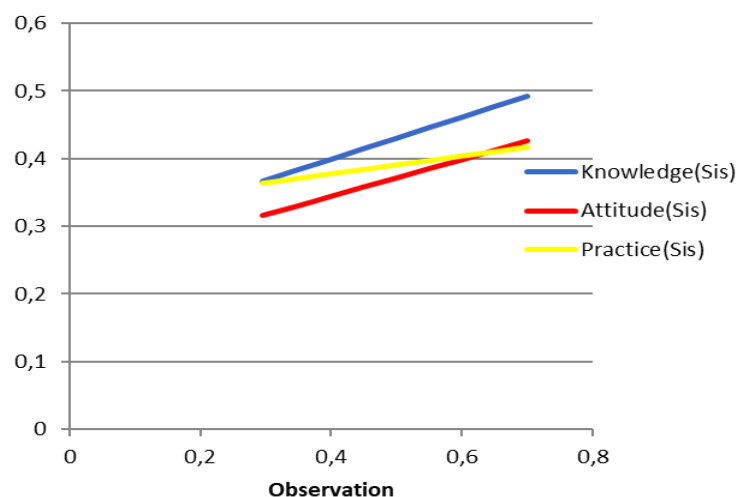
Komunitas Sekolah	Peran Agen				KAP Siswa			Rata-rata KAP
Variabel	Kepala Sekolah	Guru	Komite Sekolah	Orang Tua	K	A	P	
S	46.77	32.00	34.85	47.81	48.42	27.93	51.76	42.70
Ω_i	29.64	39.71	52.80	37.08	31.00	27.74	13.11	23.95
A_i	23.59	28.29	12.35	15.11	20.57	44.33	35.12	33.34

Perkembangan KAP siswa akibat perubahan agen komunitas sekolah, akan diberikan sebagai berikut.



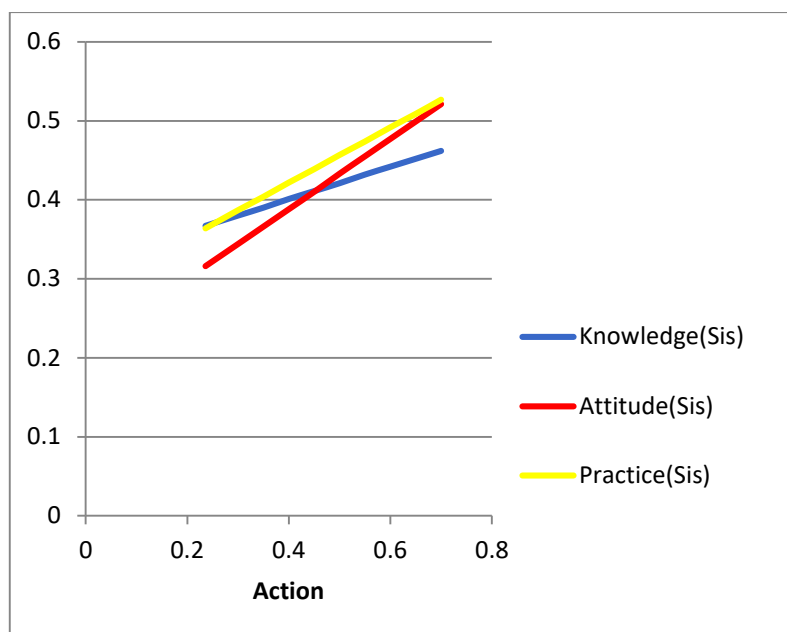
Gambar 3. Perkembangan KAP Siswa akibat Perubahan Kondisi Agen

Berdasarkan gambar diatas kemampuan praktik siswa lebih meningkat akibat adanya perubahan dari kondisi kepala sekolah.



Gambar 4. Perkembangan KAP Siswa akibat Pengamatan Perubahan Agen

Berdasarkan gambar diatas pengetahuan siswa lebih meningkat akibat adanya perubahan dari observasi kepala sekolah.



Gambar 5. Perkembangan KAP Siswa akibat Perubahan Tindakan Agen

Berdasarkan gambar diatas kemampuan praktik siswa lebih meningkat akibat adanya perubahan dari tindakan Agen. Secara keseluruhan kemampuan praktik siswa lebih baik dari pengetahuan dan sikap siswa akibat perubahan kondisi, pengamatan dan tindakan dari Agen.

5. KESIMPULAN

Penelitian ini menunjukkan bahwa Partially Observable Markov Decision Processes (POMDP) efektif untuk memahami dan mengelola perilaku pembelajaran di sekolah. Hasilnya menunjukkan peningkatan signifikan dalam pengetahuan, sikap, dan praktik (KAP) siswa.

6. DAFTAR PUSTAKA

- Boyen, X., & Koller, D. (2013). Tractable inference for complex stochastic processes. *ArXiv Preprint ArXiv:1301.7362*.
- Dean, T., Boutilier, C. T., & Hanks, S. (1999). Decision-theoretic planning: Structural assumptions and computational leverage. *Journal of Artificial Intelligence Research*, 11, 1–94.
- Doucet, A., De Freitas, N., & Gordon, N. J. (2001). *Sequential Monte Carlo methods in practice* (Vol. 1, Issue 2). Springer.
- Hauskrecht, M. (2000). Value-function approximations for partially observable Markov decision processes. *Journal of Artificial Intelligence Research*, 13, 33–94.
- Jennings, N. R. (1996). Coordination techniques for distributed artificial intelligence.

- Jensen, D., Atighetchi, M., Vincent, R., & Lesser, V. (1999). Learning quantitative knowledge for multiagent coordination. *AAAI/IAAI*, 24–31.
- Kaelbling, L. P., Littman, M. L., & Cassandra, A. R. (1998). Planning and acting in partially observable stochastic domains. *Artificial Intelligence*, 101(1–2), 99–134.
- Kanazawa, K., & Dean, T. (1989). A model for reasoning about persistence and causation. *Computational Intelligence*, 5(2), 142–150.
- Koller, D., & Lerner, U. (2001). Sampling in factored dynamic systems. In A. Doucet, N. De Freitas, & N. J. Gordon (Eds.), *Sequential Monte Carlo Methods in Practice* (pp. 445–464). Springer.
- Lesser, V., Decker, K., Wagner, T., Carver, N., Garvey, A., Horling, B., Neiman, D., Podorozhny, R., Prasad, M. N., & Raja, A. (2004). Evolution of the GPGP/TAEMS domain-independent coordination framework. *Autonomous Agents and Multi-Agent Systems*, 9(1), 87–143.
- Li, T., Choi, M., Fu, K., & Lin, L. (2019). Music sequence prediction with mixture hidden Markov models. In *2019 IEEE International Conference on Big Data (Big Data)* (pp. 6128–6132).
- Metelli, A. M., Mutti, M., & Restelli, M. (2018). Configurable Markov decision processes. In *International Conference on Machine Learning* (pp. 3491–3500).
- Mislevy, R. J., Steinberg, L. S., & Almond, R. G. (2003). Focus article: On the structure of educational assessments. *Measurement: Interdisciplinary Research and Perspectives*, 1(1), 3–62.
- Monahan, G. E. (1982). State of the art—a survey of partially observable Markov decision processes: Theory, models, and algorithms. *Management Science*, 28(1), 1–16.
- Murphy, K., & Russell, S. (2001). Rao-Blackwellised particle filtering for dynamic Bayesian networks. In A. Doucet, N. De Freitas, & N. J. Gordon (Eds.), *Sequential Monte Carlo Methods in Practice* (pp. 499–515). Springer.
- Takikawa, M., D’Ambrosio, B., & Wright, E. (2012). Real-time inference with large-scale temporal Bayes nets. *ArXiv Preprint ArXiv:1301.0603*.
- Wooldridge, M. (1999). Intelligent agents. In *Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence* (pp. 27–73).