The Thirty-Second International Florida
Artificial Intelligence Research Society Conference (FLAIRS-32)

Emergence in Multi-Agent Systems

Yan Zhao, Eugene Santos Jr.
Thayer School of Engineering at Dartmouth College
yan.zhao.th@dartmouth.edu, eugene.santos.jr@dartmouth.edu

Abstract

In a multiagent system or MAS, due to agent interactions, the agents as a group may make decisions that none of them would make alone; this phenomenon is called emergence. Emergence is characterized by an unanticipated system behavior caused by nonlinear interactions. This paper detects such emergence in a MAS by analyzing agent behaviors across two simple strategies. In the first strategy, agents make decisions based on the local information; in the second strategy, agents make decisions based on global information provided via communication. The proposed method identifies when and how nonlinear interactions cause behavior change, and quantitatively defines emergence based on the change in team performance. It then proves several theorems about emergence in a MAS. Experimental results on several benchmarks demonstrate the promising performance of the proposed framework in detecting emergence in a MAS.

Introduction

A multiagent system (MAS) is "a system with many agents who interact to reach goals" (Wu, Zilberstein, and Chen 2011). Due to agent interactions, a phenomenon called emergence could occur, which significantly impacts system reliability and predictability. We study it by comparing system behavior with and without agent interaction, finding that although interaction improves average performance, it may cause agents to perform worse in certain cases.

The contributions of this paper are threefold: 1) quantitatively defined emergence in a MAS, 2) found optimal communication strategy, and 3) analyzed risky actions.

The rest of the paper is organized as follows: the **Background** introduces emergence. Next, the **Method** details an emergence framework. Then, we proceed to the **Experiment**. Lastly, we present conclusions and future directions.

Background

This section introduces emergence and a MAS model.

Emergence has been defined as "higher-level properties resulting from lower-level causal interactions" (Mill 1884). In a MAS, agent interactions cause emergence, and agent often interact via communication. Even though researchers have heavily studied various aspects of emergence (Goldstein 1999, Bar-Yam 2004, Ryan 2007, Halley and Winkler 2008, Becker et al. 2009, Haglich et al. 2010, Santos et al. 2013, Santos and Zhao 2017, Santos et al. 2018, Zhao and Santos 2018), they have not studied how agent interaction via communication can influence emergence. Therefore, we study emergence by comparing the behavior difference of two teams, where one team communicates to make cooperative actions while the other does not communicate and make independent decisions. Their behavior difference reflects the impact of agent communication on emergence.

We model a MAS as decentralized partially observable Markov decision processes (Dec-POMDPs) (Bernstein 2000), which is formally defined as a 9-tuple $(I, S, \{A_i\}, T, R, \{\Omega_i\}, O, b^0, H)$. In a Dec-POMDPs, I is a finite set of n agents. S is a finite set of states. A_i is the set of agent i's actions. $T: S \times A \times S \to [0,1]$ is the state transition probability function. $R: S \times A \times S \to R$ is the reward function. Ω_i is the set of agent i's observations. $O: \Omega \times A \times S \to [0,1]$ is the observation probability function. b^0 is the initial state distribution named as belief state. H is the problem horizon. Agent i's observation history at t is oh_i^t, while a joint observation history is joh^t. An action history is ah^t, and a joint action history is jah^t. A joint observation action history is noted as $joah^t$. Given $joah^t$, a distribution over the current state at t is $b^t(s|joah^t)$ (Wu et al. 2011). Solving Dec-POMDPs has NEXP complexity (Bernstein et al. 2000). Researchers have proposed various optimal and approximate solvers, but extant solvers cannot solve policies for both types of agents. Therefore, we propose an approximate Dec-POMDPs solver to detect emergence.

Method

This section defines emergence in a MAS and introduces an approximated Dec-POMDPs solver to detect emergence.

Emergence Definition in a MAS

First, we define a communicative team (CT) as a team of agents who communicate with each other in decision-making processes and define a non-communicative team (NC) as a team of agents who do not communicate. CT needs an optimal joint policy π^*_{CT}, which maps joint observation histories to joint actions. NC needs an optimal local policy $\pi^*_{NC} = \langle \pi^*_1, \ldots, \pi^*_n \rangle$, which maps local observation histories to individual actions. Then, we define several terms for emergence definition as follows:

Def 3.1 Define $path = \langle joh^t, jah^t \rangle$ as a joint observation action history at step $t = H - 1$, and define $PATH(\pi)$ as the set of paths induced from policy π.

Def 3.2 Define a pair of paths, $path' = \langle joh^{t\prime}, jah^{t\prime} \rangle \in PATH(\pi^*_{CT})$ and $path'' = \langle joh^{t\prime\prime}, jah^{t\prime\prime} \rangle \in PATH(\pi^*_{NC})$ as partially matched paths iff: $joh^{t\prime} = joh^{t\prime\prime}$ and $jah^{t\prime} \neq jah^{t\prime\prime}$;

We use partially matched paths to evaluate agent behavior change measured by accumulated rewards (ARs), where AR is the sum of rewards of all steps.

Def 3.3 Define $f(pt)$ as the CDF of ARs of $pt, pt = \{path', path''\}$.

Next, we compare the CDFs with Quantile-Quantile (QQ) plots for $path'$ and $path''$, and perform linear regression on the QQ plot, which yields a linear function $y = ax + b + \epsilon$ and R^2.

Def 3.4 Define **Type-I emergence** if $R^2 > 0.95$ & $a \neq 1$.

A slope of $a \neq 1$ means the length of $f(path'')$'s domain differs from the length of $f(path')$'s domain. Therefore, communication changes AR distribution. A correlation coefficient larger than 0.95 means each NC AR maps to a higher CT AR linearly. Therefore, communication does not change the relative order of possible rewards agent can earn.

Def 3.5 Define **Type-II emergence** if $R^2 \leq 0.95$.

Type-II emergence reflects a nonlinear change of AR distribution. CT may have a different number of ARs than NC does. The relative order of ARs in CT may differ from that of NC's. The lower bound of AR distribution of CT may be smaller than that of NC's. Such changes are more drastic than those in **Type-I emergence.**

Theoretical Foundations of Emergence in a MAS

This subsection proves three theorems on emergence. **Theorem 1** proves that CT has higher or equal expected AR than NC does. **Theorem 2** proves that if CT only communicates at $H - 1$, CT earns higher or equal AR than NC does. **Theorem 3** proves that if CT communicates at $t < H - 1$, CT may get lower expected AR than NC does.

Table 1. MCGD for CT

1	for each step $t \in [0, H-1]$
2	for each joint observation joh^t
3	evaluate every joint action $\vec{a}^t_1, \ldots, \vec{a}^t_k \in A$
4	if $s(\vec{a}^t_j) \gg s(\vec{a}^t_r), r \neq j$ outperforms, pick $\vec{a}^t_j$
5	else if $s(\vec{a}^t_{k1}) \approx \cdots \approx s(\vec{a}^t_{km})$, add $\vec{a}^t_{k1} \ldots \vec{a}^t_{km}$ to PL
6	and randomly pick from $\vec{a}^t_{k1} \ldots \vec{a}^t_{km}$
7	find the best joint actions in PL via gradient descent
8	repeat step 1-7, pick the best joint policy π^*_{CT}

$$S(\vec{a}^j_i) = \textstyle\sum_{s \in S} b^j(s|joah^j) Q(s, \vec{a}^j_i, j) \quad (1)$$

$$\textstyle\sum_{\vec{s} \in \vec{S}} b^0(s^0) \prod_{t=0}^{H-1} T(s^t, \vec{a}^t, s^{t+1}) \prod_{t=0}^{H-2} O(\vec{o}^{t+1}, \vec{a}^t, s^{t+1}) \quad (2)$$

$$\textstyle\sum_{\vec{s} \in \vec{S}} [P(\vec{s}) \sum_{t=0}^{H-1} R(s^t, \vec{a}^t, s^{t+1})] \quad (3)$$

$$\textstyle\sum_{k=1}^{3} P(pt_k) AR(pt_k) \,/\, \sum_{k=1}^{3} P(pt_k) \quad (4)$$

Theorem 1 $E(AR(\pi^*_{CT})) \geq E(AR(\pi^*_{NC}))$.

Proof: We can build a π'_{CT} from π^*_{NC}: at $t \in [0, H)$, add $\pi'_{CT}(joh^t) = \vec{a}^t$, where $\pi^*_i(oh^t_i) = a^t_i$ for agent i, $joh^t = \langle oh^t_1, \ldots, oh^t_n \rangle$, and $\vec{a}^t = \langle a^t_1, \ldots, a^t_n \rangle$. Therefore, $E(AR(\pi^*_{CT})) = E(AR(\pi^*_{NC}))$. In addition, if $\exists \vec{a} \in A$ and $joh^t, t \in [0, H)$ s.t. $R(joh^t, \vec{a}) > R(joh^t, \vec{a}^t)$, where $R(joh^t, \vec{a}) = \sum_{s \in S} b^t(s|joah^t) \cdot \sum_{s' \in S} R(s, \vec{a}, s') \cdot T(s, \vec{a}, s')$, then replacing $\vec{a}^t$ with $\vec{a}$ improves $AR(\pi'_{CT})$. In this case, $E(AR(\pi^*_{CT})) > E(AR(\pi^*_{NC}))$. QED.

Theorem 2 If $\langle path', path'' \rangle$ only differ at $\vec{a}^{H-1}$, then $E(ARs(path')) \geq E(ARs(path''))$.

Proof: Both paths have the same $b^{H-1}(s)$, and $\vec{a}^{H-1}_{NC} = \langle \pi^*_1(oh^{H-1}_1), \ldots, \pi^*_n(oh^{H-1}_n) \rangle$, $\vec{a}^{H-1}_{CT} = \pi^*_{CT}(joh^{H-1})$. If $\sum_{s \in S} b^{H-1}(s) R(s, \vec{a}^{H-1}_{NC}) > \sum_{s \in S} b^{H-1}(s) R(s, \vec{a}^{H-1}_{CT})$, then mapping joh^{H-1} to $\vec{a}^{H-1}_{NC}$ will increase $E(AR(\pi^*_{CT}))$, which contradicts the definition of π^*_{CT}. Therefore, $\sum_{s \in S} b^{H-1}(s) R(s, \vec{a}^{H-1}_{NC}) \leq \sum_{s \in S} b^{H-1}(s) R(s, \vec{a}^{H-1}_{CT})$, and $E(ARs(path'')) \leq E(ARs(path'))$. QED.

Theorem 3 If $\langle path', path'' \rangle$ differ at step $t \in [0, H-2]$, then $E(ARs(path')) \geq or \leq E(ARs(path''))$.

Proof: Assume that they only differ at $t \in [0, H-2]$, $(\vec{a}^t_{CT}, \vec{a}^t_{NC})$. There are $K = |\Omega|^{H-1-t}$ paths with $\vec{a}^t_{CT}$, $\{p_k | k = 1, \ldots, K\}$, and K paths with $\vec{a}^t_{NC}$. Per **Theorem 1**, $\sum_k Pr(p_k) \cdot E(AR(p_k)) \geq \sum_k Pr(p'_k) \cdot E(AR(p'_k))$. Therefore, it is possible that $E(AR(p_{k1})) > E(AR(p'_{k1}))$ when $Pr(p_{k1}) > Pr(p'_{k1})$ and $E(AR(p_{k2})) < E(AR(p'_{k2}))$ otherwise. If $path' = p_{k1}, and\ path'' = p'_{k1}$, then $E(ARs(path')) > E(ARs(path''))$; if $path' = p_{k2}, and\ path'' = p'_{k2}$, then $E(ARs(path')) < E(ARs(path''))$. QED.

An Approximate Dec-POMDPs Solver: MCGD

This subsection proposes a solver called Monte Carlo sampling and gradient descent or MCGD. Table 1 shows its pseudocode for CT. First, it samples paths (1-5). For a path, at $j \in [t, H)$, it computes $b^t(s|joah^t)$. Then it samples a state s^j. Next, if $j = t$, it picks $\vec{a}^t_i$; otherwise, it scores each

$\vec{a}_i^j, i \in [1, |A|]$ per (1). Later, it ranks all actions and picks one. Then it samples a joint observation.

When a path pt finishes, it computes its probability $P(pt)$ and expected $AR(pt)$ per (2) and (3). After it samples enough paths, it picks the top three paths in Hp and uses their weighted average score as $\vec{a}_i^t$'s score per (4). If $\vec{a}_k^t$ is significantly better, it maps joh^t to $\vec{a}_k^t$; otherwise, it puts comparable actions $\{\vec{a}_{k1}^t, \vec{a}_{k2}^t, \dots, \vec{a}_{km}^t\}$ in a pool PL. After it computes joint actions for all joh^{H-1}, it applies gradient descent to select the best combination of joint actions in PL (6-7). It repeats this process several times and chooses the best policy as π_{CT}^*. The procedure for NC is similar.

Experiment

This section first introduces three Dec-POMDPs benchmarks. It then compares MCGD with a state-of-the-art exact solver: generalized multiagent A* with incremental clustering and expansion (GMAA*-ICE) (Oliehoek, F. A. et al. 2013). Lastly, it visualizes and explains emergence.

Benchmarks and Evaluation of MCGD

The first benchmark is Dec-Tiger (DT) (Nair et al. 2003): There are two doors, behind one is a tiger and behind the other is treasure. Two agents look for treasure via three actions: Listen, Open Left, and Open Right. An agent hears the tiger with a 0.85 probability. If both agents open the treasure door, they get a reward of 20. If both open the tiger door, they get -50. If one opens the treasure door, they get 9. If one opens the tiger door, they get -101. If they open two doors, they get -100. The game resets once a door opens.

The second benchmark is Two Generals (2G) (http://www.fransoliehoek.net). Two generals can attack the enemy or observe its status (small or large) with an 85% accuracy. If the enemy is small and they attack simultaneously, they get a reward of 5; if the enemy is large and they attack simultaneously, they get -20; if one attacks, they get -10. The game resets once an attack occurs.

The third benchmark is Recycling Robots (RR) (Sutton and Barto 1998). Two robots pick up cans of two sizes. A robot's battery can be high and low, which is observable to itself with a 0.9 accuracy. A robot can recharge, pick up a small can by itself with reward 2 and jointly pick up a large can with reward 5. Every time a robot picks up a can, the battery becomes low with a 0.3 probability for a small can and with a 0.5 probability for a large can. If the battery is low and a robot picks a small or large can, it drains the battery with a 0.2 or 0.3 probability and gets recharged. Although there are only two agents in these problems, their solutions are easily applicable to problems with more agents. We apply both MCGD and GMAA*-ICE to solve them on an Intel Xeon@2.66GHz CPU with 512GB ram. Table 2 lists policy scores for NC and CT at various horizons by MCGD, which is at least 98% of the optimal score. Since GMAA*-ICE runs out of memory (512GB) for DT at H=6, 2G at H=9 and RR at H=5, we only show results where both algorithms can finish. Type I and Type II show that MCGD detects at least 95% emergence cases.

Table 2. Policy Score and Emergence Case by MCGD

D	DT			2G						RR	
H	3	4	5	3	4	5	6	7	8	3	4
NC	5.2	4.8	7	-2.9	-2.4	-3.2	-3.6	-7.3	-8.6	10	12.4
CT	13	23	26.8	-0.4	-0.6	-1.2	-1.8	-2.5	-3.3	10.2	12.9
Type I	0	36	116	4	13	60	289	1304	3824	4	11
Type II	8	16	112	0	4	41	168	795	3310	8	28

Table 3. Emergence Case Based on AR Relation

Case	1	2	3	4
E(CT AR)>E(NC AR)	Y	Y	N	N
lb(CT AR)>lb(NC AR)	Y	N	Y	N

In different emergent path pairs, the relationship between CT and NC scores are classified into four cases (Table 3). Case 1 indicates that without communication, two agents will take uncoordinated actions and that communication improves not only expected AR but also the lower bound of AR distributions. Case 2 means although CT gets higher expected AR, CT can earn lower rewards. Case 3 and Case 4 shows that CT gets worse results than NC.

Error! Reference source not found. shows the frequencies of emergence cases in the benchmarks. In all benchmarks, as horizon increases, more case 3 and 4 path pairs occur, as suggested by **Theorem 3**. However, different benchmarks have different leading emergence case and relative ranking of different cases. In general, case 3 and 4 are less likely than case 1 and 2, as suggested by **Theorem 1**.

Visualization and Explanation of Emergence

Error! Reference source not found. shows the QQ-plots of 2G at H=5. The majority part of QQ-plots is above the baseline, indicating that CT rewards are higher than NC (**Theorem 1**). There are several types of QQ plots. In the first type (case 1 of **Type II emergence**) such as the purple one, CT has higher expected AR (**Theorem 1)** and no worse AR than NC in each state (**Theorem 2)**. In the second type (case 2 of **Type II**) such as the yellow, red, and black one, CT has higher expected and maximal AR, but lower minimal AR (**Theorem 2&3**). In other words, CT chooses joint actions that give better expected AR while introducing the possibility of worse minimal reward. In the third type (case 3&4 of **Type II**) such as the green and blue one, CT has lower expected AR and lower minimal AR than NC (**Theorem 3**), which suggests CT chooses inferior joint actions in this path to maximize rewards in paths with higher probabilities.

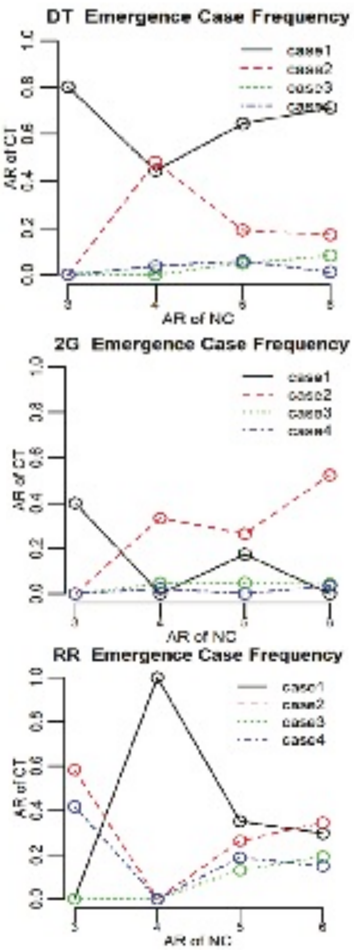

Figure 1. Emergence Case Frequency vs Problem Horizon

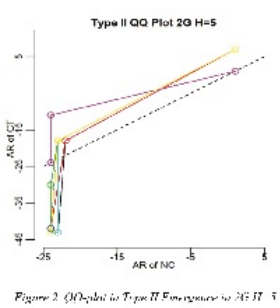

Figure 2. QQ-plot in Type II Emergence in 2G H=5

Conclusion

This paper proposed a quantitative emergence detection framework based on the linearity of reward distributions. It also proved several theorems about emergence in a MAS and explored several features that lead to emergence. Experimental results on several MAS planning problems showed that this framework could detect emergence by efficiently comparing two teams' policies. To the best of the authors' knowledge, this framework is the first of its kind for emergence detection.

References

Bar-Yam, Y. 2004. A Mathematical Theory of Strong Emergence Using Multiscale Variety. *Complexity* 9(6): 15–24.

Becker, R. et al. 2009. Analyzing Myopic Approaches for Multi-Agent Communication. *Computational Intelligence 25*(1): 31-50.

Bernstein, D. S., Zilberstein, S. and Immerman, N., 2000. The Complexity of Decentralized Control of Markov Decision Processes. In *Proceedings of the Sixteenth Conference on Uncertainty of artificial intelligence,* 32–37. Morgan Kaufmann Publishers Inc.

Goldstein, J. 1999. Emergence as A Construct: History and Issues. *Emergence* 1(1): 49-72.

Haglich, P., Rouff, C., and Pullum, L. 2010. Detecting Emergence in Social Networks. *In the Second International Conference on Social Computing,* 693-696. IEEE

Halley, J., and Winkler, D. A. 2008. Classification of Emergence and Its Relation to Self-Organization. *Complexity* 13(5): 10–15.

Mill, J. S. 1884. A System of Logic, Ratiocinative and Inductive: Being a Connected View of the Principles of Evidence and the Methods of Scientific Investigation Vol. 1. Longmans, green, and Company.

Nair, R. et al. 2003. Taming Decentralized POMDPs: Towards Efficient Policy Computation for Multiagent Settings. In *Proceedings of the Eighteenth International Joint Conference on Artificial Intelligence*, 705-711. Acapulco, Mexico.: International Joint Conferences on Artificial Intelligence, Inc.

Ryan, A. J. (2007). Emergence is Coupled to Scope, Not Level. *Complexity* 13(2): 67–77.

Oliehoek, F. A. et al. 2013. Incremental Clustering and Expansion for Faster Optimal Planning in Decentralized POMDPs. *Journal of Artificial Intelligence Research* 46: 449–509.

Santos E. E. et al. 2013. Modeling emergent border-crossing behaviors during pandemics. In *Proceedings of the 2013 SPIE Defense, Security, and Sensing Conference,* 87110Z. Baltimore, Maryland.

Santos E. E. et al. 2018. Modeling Social Resilience in Communities. in *IEEE Transactions on Computational Social Systems*, 5(1): 186-199.

Santos, E., and Zhao, Y. 2017. Automatic Emergence Detection in Complex Systems. *Complexity* 2017. 24 pages.

Seuken, S., and Zilberstein, S. 2007. Memory-Bounded Dynamic Programming for DEC-POMDPs. In *Proceedings of the 20th International Joint Conference on Artificial Intelligence*, 2009-2015. Hyderabad, India.: International Joint Conferences on Artificial Intelligence, Inc.

Sutton, R. S., and Barto, A. G. 1998. *Introduction to reinforcement learning*. Cambridge.: MIT press.

Szer, D., and Charpillet, F. 2005. An Optimal Best-First Search Algorithm for Solving Infinite Horizon DEC-POMDPs. In *Proceedings of the Sixteenth European Conference on Machine Learning 389-399. Springer,* Berlin, Heidelberg.

Zhao, Y. and Santos, E. 2018, A Failure of Collective Intelligence. In *Proceedings of the 2018 IEEE/WIC/ACM International Conference on Web Intelligence (WI)* 361-366. Santiago, Chile.