Intertemporal Equilibrium in Online Routing Games*

László Z. Varga

Faculty of Informatics, ELTE Eötvös Loránd University, H-1117 Budapest, Hungary
http://people.inf.elte.hu/lzvarga

Abstract. In order to ensure global behaviour of decentralized multi-agent systems, we have to have a clear understanding of the issue of equilibrium over time. The evolutionary dynamics within a game is investigated in online routing games. The progress beyond the state-of-the-art is that we introduce the notion of intertemporal equilibrium in the study of the global behaviour of decentralized multi-agent systems, we define quantitative values to measure the intertemporal equilibrium, we use these quantitative values to evaluate a realistic scenario, and we give an insight into the influence of the designed intertemporal expectations of the agents on the global behaviour of decentralized multi-agent systems. An interesting result is that if the multi-agent system is designed in a way that agents have less precise knowledge of the future, then it leads to better global behaviour of the multi-agent system.

Keywords: Control of the Global Behaviour of Decentralized MAS · Ensuring Design Goals · Agent Theories and Models.

1 Introduction

In order to be able to define and measure design criteria, designers need formal models. Currently the best model of multi-agent decision making is based on game theory. The designers prefer multi-agent systems with an equilibrium, because the equilibrium seems to be a stable state of the system. If the equilibrium meets the design criteria, then we can ensure that the multi-agent system behaves in accordance with the design goals. The classic game theory models assume an idealistic situation: all the agents know what the equilibrium is, all the agents know what other agents do, and all the agents know what their role is in the equilibrium. The agent behaviour goes in cycles: the agents perceive their environment, decide what action to perform, and then perform the action. Can we ensure that multi-agent systems go to the equilibrium through these feedback cycles and stay in the equilibrium, as intended by the designers?

Many real world applications are *continuously evolving games*: agents join the game in a sequence, they influence the game for a while, and then they quit the

* This work was carried out in the project EFOP-3.6.3-VEKOP-16-2017-00001: Talent Management in Autonomous Vehicle Control Technologies The Project is supported by the Hungarian Government and co-financed by the European Social Fund.

game. In these games, the decisions of the agents are often *intertemporal choices*: the current decision of the agent may affect the utility of the agents in the future. The equilibrium of such evolving games is the *intertemporal equilibrium*.

Intertemporal equilibrium [6] has two interpretations in economic theory. One interpretation is related to the intertemporal aspect of the choice. The other interpretation is related to the temporal aspect of the equilibrium: at any given time, the economy is in disequilibrium, and the equilibrium can be interpreted only in the long term. In this paper, we focus on the latter interpretation, and we take into account that agents have intertemporal expectations.

In order to study how to control global behaviour over time, we take the large-scale and open multi-agent system of the road traffic application area, and in particular the *online routing problem*. The online routing problem is a network with traffic flows going from a source node to a destination node. The agents of the traffic flows continuously enter the network at the sources, they choose a full route to the destination of their trip, and quit the network at the destination. The traffic is routed in a congestion sensitive manner.

2 Related Work

The *static equilibrium* is an important concept of game theory. Algorithmic game theory [8] investigated the routing problem where decentralised autonomous decision making is applied by the traffic flows. This game theory model is in line with the assumption of the traffic engineers, who assume that the traffic is always assigned in accordance with the static equilibrium [15]. The potential function is used to prove the existence of equilibria, and an upper limit on the price of anarchy is also proved [8].

The *evolutionary dynamics* of games is usually investigated in repeated games where the agents receive feedback by observing their own and other agents' actions and utility, and in the next game they may change their own actions. The potential function method is extended to prove that the repeated routing game converges to the static equilibrium [5]. Another type of feedback is *regret minimisation*, where agents compare their actually experienced utility with the best possible utility in retrospect. It is proved that if the agents of the routing game select actions to minimize their regret, then their behaviour converges to the static equilibrium [1]. The repeated game approach captures the evolutionary dynamic between routing games, but not within the routing game.

The *deterministic queuing model* is an approximation to investigate how traffic flows evolve over time. The Nash flows over time in non-atomic queuing networks is characterised and several bounds on the price of anarchy are proved in [7]. Single source fluid queuing networks reach a steady state in finite time if the inflow does not exceed the capacity of the network [4]. The queuing model does not have usage dependent cost of the edges if the inflow is below the maximum capacity, because in this flow range the edge has a constant delay.

The evolutionary dynamic inside the routing game is captured by the *online routing game* model, where the traffic flow is made up of individual agents who follow each other, and the agents of the traffic flow decide individually on their actions based on the real-time situation. The reader is referred to the openly accessible article [9] for the formal description of the online routing game model.

It is proved [10] that if the agents of the online routing game try to maximise their utility computed from the real-time situation (without taking into account any expectation), then equilibrium is not guaranteed, although a static equilibrium exists. In order to facilitate the agents to make predictions and include future conditions in their decisions, intention-aware prediction methods were proposed. In the intention-aware prediction methods, the agents communicate their intentions to a service. The service aggregates the data about the agent collective, and it sends a feedback to the agents [2]. The *intention-aware* [16] and the *intention propagation* [3] approaches are based on this scheme.

It is proved [9] that there is no guarantee on the equilibrium, even if intention-aware prediction is applied. However, it is proved [11] that in a small but complex enough network of the Braess paradox, the agents might just slightly be worse off in the worst case with real-time data and prediction. It is also proved [13] that in the network of [11], the system converges to the static equilibrium within a relatively small threshold. The conjecture in [12] says that the system converges to the static equilibrium in bigger networks as well if simultaneous decision making is prevented. This conjecture neither has been proved nor refuted analytically.

3 Intention-Aware Prediction Methods

The formal description of the algorithms of two intention-aware prediction methods were presented at [14]: the detailed prediction method and the simple prediction method.

The *detailed prediction method* takes into account all the intentions already submitted to the service, then it computes what will happen in the future if the agents execute the plans assigned by these intentions, and then it computes for each route in the network the predicted travel time by taking into account the predicted future travel times for each road of the route. The prediction algorithm used in [16] is close to this detailed prediction method, but the main difference is that the prediction algorithm of [16] uses probabilistic values, while the detailed prediction method is deterministic.

The *simple prediction method* also takes into account all the intentions already submitted to the service, however when it computes for each route in the network the predicted travel time, then it takes into account only that travel time prediction for each road which was computed at the last intention submission. This way, the simple prediction method needs a little bit less computation. The simple prediction method is a kind of approximation and does not try to be an exact prediction of the future. As time goes by, if no new prediction is generated for a road, then the simple prediction method "evaporates" the last prediction for that road, like the bio-inspired technique of [3].

The investigations in [14] left the issue of the convergence to the equilibrium open.

4 Experimental Set-up

In order to investigate empirically the intertemporal equilibrium, a region of Budapest (shown in Fig. 1) was modelled in the simulation software of [10]. The figure shows the route choices towards the destination *Rákóczi* bridge (E in the figure) from two sources: the suburban area (A) and the intercity road (B). Information on the traffic flow going on these roads can be obtained from the web site[1] of the Hungarian Public Road Non-profit PLC. The minimum travel time in minutes (fixed part of the cost function) for the roads is 1.5 times the distance. The variable part of the travel time is $roadlength * flow \div 10$, thus the cost functions are the following: $c_{(A,C)}(flow) = 2.1 + 1.4 * flow \div 10$, $c_{(B,C)}(flow) = 1.5 + 1.0 * flow \div 10$, $c_{(C,D)_{north}}(flow) = 6.0 + 4.0 * flow \div 10$, $c_{(C,D)_{south}}(flow) = 10.2 + 6.8 * flow \div 10$, $c_{(D,E)}(flow) = 1.8 + 1.2 * flow \div 10$, where the cost is in $minute$ and the traffic flow is in $car \div minute$.

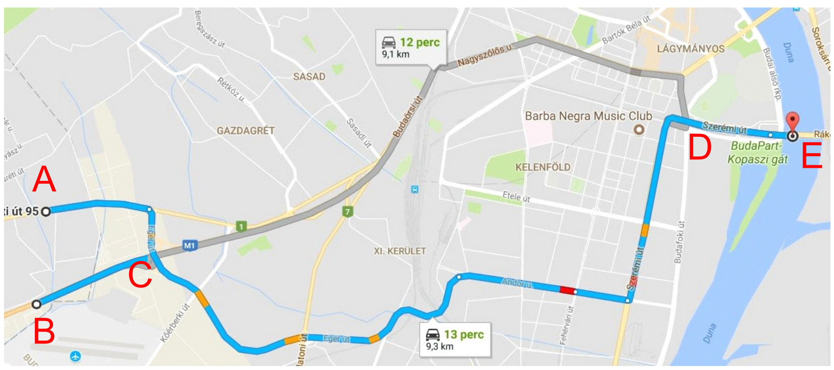

Fig. 1. The Google Map extract showing the realistic scenario of the experiments

The experiment simulates a 90 minute long rush hour period extended with a 17 minute initial period to populate the roads to some extent. Several experiments were run at traffic flow values from 2.5 $car \div minute$ to 30 $car \div minute$ in steps of 2.5. The incoming flow values were the same at points A and B. Simultaneous decision making was excluded.

All the experiments were executed in three versions using three different routing strategies: 1) no prediction routing strategy, 2) detailed prediction routing strategy, and 3) simple prediction routing strategy. The no prediction routing strategy is the simple naive (SN) online routing game of [10], where the routing strategy selects the shortest travel time observable in the real-time status of the network. The latter two strategies are intention-aware routing strategies where the routing strategy selects the shortest predicted travel time using the corresponding prediction method as described in Section 3.

[1] http://internet.kozut.hu/Lapok/forgalomszamlalas.aspx

5 Measure of Intertemporal Equilibrium

In order to explain how we quantify the intertemporal equilibrium, we measured the travel times of the trip $A - E$ with all three routing strategies at flow value 20 $car \div minute$. Due to lack of space, we cannot show the measurements here. The time period of a bit more than 450 minutes was selected to show that the travel times do not seem to converge to a steady value. This is in line with our interpretation of intertemporal equilibrium: at any given moment the system is in disequilibrium and the equilibrium can be interpreted only in a time period.

We can also see from the measurements that the travel time seems to remain within a limit from a kind of equilibrium value. The worst case difference might be big, but if the system fluctuates around a kind of equilibrium, then we consider it as intertemporal equilibrium.

We can also see from the measurements that sometimes there are big differences in the travel times of the agents that ended their trips at almost the same elapsed time of the experiment. This means that there was an agent which arrived through a non congested route, and another agent arrived through a congested route almost at the same time. The smaller travel time differences seem to coincide with smaller swings in the system. The smaller swings are probably because the agents have almost equal intertemporal choices during the experiment, although the system is fluctuating all the time. We call this phenomenon "quasi equilibrium within the disequilibrium".

Based on the above observations, we define the quantitative measure of intertemporal equilibrium the following way:

Definition 1. *Let* $ORG =< t, T, G, c, r, k >$ *be an online routing game over the finite sequence of time steps* t*. Let* $c_{r_i}(\tau)$ *be the cost of the agent of trip* $r_i \in r$ *when it exits the game at time step* $\tau \in t$*. Let* e_{r_i} *be the static equilibrium travel time for trip* r_i*. The* `measure of intertemporal equilibrium` *of* ORG *is* $< WD, AD, QE >$ *where* $WD = \max_{r_i \in r}(\max_{\tau \in t}((c_{r_i}(\tau) - e_{r_i}) \div e_{r_i}))$ *,* $AD = \text{avg}_{r_i \in r}((\text{avg}_{\tau \in t}(c_{r_i}(\tau) - e_{r_i})) \div e_{r_i})$ *, and* $QE = \text{avg}_{r_i \in r}(\text{avg}_{\tau \in t}(|c_{r_i}(\tau) - c_{r_i}(\tau + 1)| \div e_{r_i}))$*.*

In the ideal case the system stays continuously in the static equilibrium, in which case the intertemporal equilibrium is $< WD = 0, AD = 0, QE = 0 >$. If $QE = 0$, then the system is in a kind of quasi equilibrium, and the WD and AD values indicate how worse this equilibrium is than the static equilibrium.

Note that the QE is an important part of the intertemporal equilibrium, and it includes much more information than usual statistical values like for example the standard deviation of travel times. Two experiments may have the same standard deviation, but they may have different QE values.

6 Evaluation of the Experiments

Before the experiments, our expectation was that we could confirm the following hypotheses:

H1: The intention-aware routing strategies produce better intertemporal equilibrium values than the non predictive routing strategy.

H2: The detailed prediction routing strategy produces better intertemporal equilibrium values than the simple prediction routing strategy, because the simple prediction method does not try to be precise.

The worst case difference values (WD) of the intertemporal equilibrium of the experiments are shown in Fig. 2. The horizontal axis is the traffic flow rate value of each experiment, in $car \div minutes$. The WD is zero at low traffic flows, because all the traffic can go on the shortest route, and this is the static equilibrium as well. At higher traffic values, the WD increases, it even reaches 1 in the case of the no prediction routing strategy. The WD of the detailed prediction routing is better than the WD of the no prediction routing in most of the experiments. The WD of the simple prediction routing is better than the WD of the detailed prediction routing. The WD values confirm hypothesis H1 in most of the experiments, but they refute hypothesis H2.

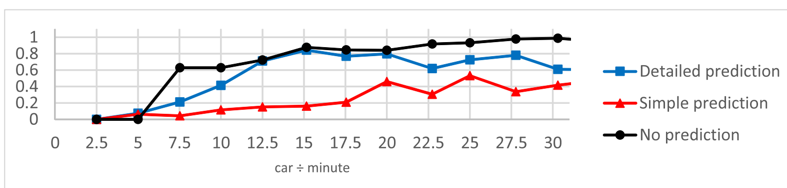

Fig. 2. The worst case difference values (WD) in the experiments

The average difference values (AD) of the intertemporal equilibrium of the experiments are shown in Fig. 3. The AD is zero at low traffic flows, like in the case of the WD. At higher traffic values, the AD increases, but it is considerably less than the WD. The AD of the detailed prediction routing is better than the AD of the no prediction routing in most of the experiments. The AD of the simple prediction routing is better than the AD of the detailed prediction routing. The AD values confirm hypothesis H1 in most of the experiments, but they refute hypothesis H2.

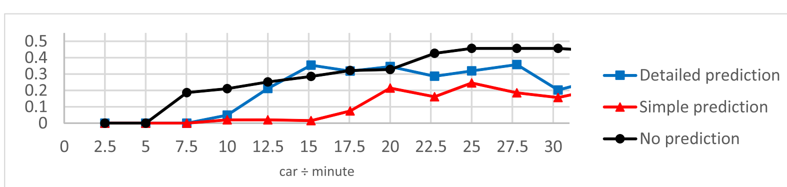

Fig. 3. The average difference values (AD) in the experiments

The quasi equilibrium values (QE) of the intertemporal equilibrium of the experiments are shown in Fig. 4. The QE is zero at low traffic flows. At higher traffic values, the QE increases, but they are much smaller for the detailed prediction routing and the simple prediction routing than for the no prediction routing. The QE of the simple prediction routing is better than the QE of the detailed prediction routing. The QE of the simple prediction routing is very close to zero. The QE values confirm hypothesis H1, but they refute hypothesis H2.

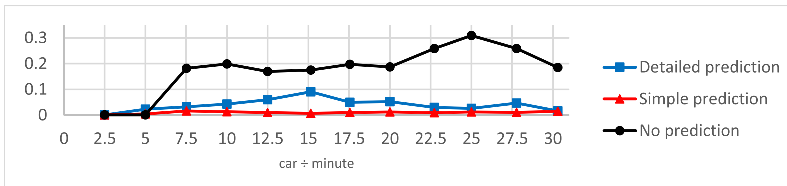

Fig. 4. The quasi equilibrium values (QE) in the experiments

7 Discussion

We have investigated the evolutionary dynamics inside games, in particular online routing games. There is a conjecture that online routing games with specific properties converge to the static equilibrium. In this paper we took a different approach to the issue of convergence to the equilibrium. Instead of proving the convergence, we studied the nature of the kind of equilibrium that seems to appear in online routing games. With this work we contribute to the better measurement and control of the global behaviour of multi-agent systems.

One of the results is that if the agents of the online routing game base their decisions only on the current situation, then the global behaviour over time is worse than in the case when they have a prediction of the future.

Another result is that knowing a more precise prediction of the future does not lead to a better global behaviour over time in the experiments. This is unexpected. In our view, this is an important new result, because it demonstrates in a controllable experiment that the selfish adaptation of the individual agents to a less precisely expected future leads to better agent system behaviour. Better knowledge of the future may not be better for the multi-agent system. This is an important design guideline for engineering multi-agent systems.

Finally, the formal definition of the measure of intertemporal equilibrium in online routing games is an important result, because it gives better insight into the dynamic behaviour of multi-agent systems. We have defined a measurement value for the quasi equilibrium in the disequilibrium. The presented experiments show that better QE values correspond to better system behaviour. This result is a guidance for the design of better multi-agent systems.

References

1. Blum, A., Even-Dar, E., Ligett, K.: Routing without regret: On convergence to nash equilibria of regret-minimizing algorithms in routing games. In: Proc. PODC '06. pp. 45–52. ACM, New York, NY, USA (2006). https://doi.org/10.1145/1146381.1146392
2. Claes, R., Holvoet, T.: Traffic coordination using aggregation-based traffic predictions. IEEE Intelligent Systems **29**(4), 96–100 (July 2014). https://doi.org/10.1109/MIS.2014.73
3. Claes, R., Holvoet, T., Weyns, D.: A decentralized approach for anticipatory vehicle routing using delegate multi-agent systems. IEEE Transactions on Intelligent Transportation Systems **12**(2), 364–373 (2011). https://doi.org/10.1109/TITS.2011.2105867
4. Cominetti, R., Correa, J., Olver, N.: Long term behavior of dynamic equilibria in fluid queuing networks. In: IPCO 2017, June 26-28, 2017. LNCS, vol. 10328, pp. 161–172 (2017). https://doi.org/10.1007/978-3-319-59250-3_14
5. Fischer, S., Vöcking, B.: On the evolution of selfish routing. In: Proc. ESA '04. LNCS, vol. 3221, pp. 323–334 (2004). https://doi.org/10.1007/978-3-540-30140-0_30
6. Hicks, J.R.: Value and Capital: An Inquiry into some Fundamental Principles of Economic Theory. Oxford University Press, 2nd edn. (1975)
7. Koch, R., Skutella, M.: Nash equilibria and the price of anarchy for flows over time. Theory of Computing Systems **49**(1), 71–97 (Jul 2011). https://doi.org/10.1007/s00224-010-9299-y
8. Nisan, N., Roughgarden, T., Tardos, E., Vazirani, V.V.: Algorithmic Game Theory. Cambridge University Press, New York, NY, USA (2007). https://doi.org/10.1017/CBO9780511800481
9. Varga, L.: On intention-propagation-based prediction in autonomously self-adapting navigation. Scalable Computing: Practice and Experience **16**(3), 221–232 (2015), `http://www.scpe.org/index.php/scpe/article/view/1098`
10. Varga, L.Z.: Online routing games and the benefit of online data. In: ATT 2014 8th International Workshop on Agents in Traffic and Transportation. pp. 88–95 (2014), `http://www.ia.urjc.es/ATT/documents/ATT2014proceedings.pdf`
11. Varga, L.Z.: Benefit of online real-time data in the braess paradox with anticipatory routing. In: IEEE International Conference on Autonomic Computing, Würzburg, Germany. pp. 245–250 (2016). https://doi.org/10.1109/ICAC.2016.68
12. Varga, L.Z.: How good is predictive routing in the online version of the braess paradox? Frontiers in Artificial Intelligence and Applications, vol. 285, pp. 1696–1697. IOS Press (2016). https://doi.org/10.3233/978-1-61499-672-9-1696
13. Varga, L.Z.: Equilibrium with predictive routeing in the online version of the braess paradox. IET Software **11**(4), 165–170 (Aug 2017)
14. Varga, L.Z.: Two prediction methods for intention-aware online routing games. In: Proc. of 5th International Conference on Agreement Technologies, Évry, France, 14-15 December 2017. pp. paper no.6. in the pre–proceedings. Springer-Verlag (2017)
15. Wardrop, J.G.: Some theoretical aspects of road traffic research. Proceedings of the Institution of Civil Engineers, Part II **1**(36), 352–378 (1952)
16. de Weerdt, M.M., Stein, S., Gerding, E.H., Robu, V., Jennings, N.R.: Intention-aware routing of electric vehicles. IEEE Transactions on Intelligent Transportation Systems **17**(5), 1472–1482 (May 2016). https://doi.org/10.1109/TITS.2015.2506900