Empirical game theory of pedestrian interaction for autonomous vehicles

Fanta Camara[1,2], Richard Romano[1], Gustav Markkula[1], Ruth Madigan[1], Natasha Merat[1] and Charles Fox[1,2,3]

[1] **Institute for Transport Studies (ITS), University of Leeds**

[2] **School of Computer Science, University of Lincoln**

[3] **Ibex Automation Ltd**

Abstract Autonomous vehicles (AV's) are appearing on roads, based on standard robotic mapping and navigation algorithms. However their ability to interact with other road-users is much less well understood. If AVs are programmed to stop every time another road user obstructs them, then other road users simply learn that they can take priority at every interaction, and the AV will make little or no progress. This issue is especially important in the case of a pedestrian crossing the road in front of the AV. The present methods paper expands the sequential chicken model introduced in (Fox et al., 2018), using empirical data to measure behavior of humans in a controlled plus-maze experiment, and showing how such data can be used to infer parameters of the model via a Gaussian Process. This providing a more realistic, empirical understanding of the human factors intelligence required by future autonomous vehicles.

1. Introduction

Autonomous vehicle technology is currently claimed by many commercial organizations involved in operating experiments which apply robotic navigation algorithms to on-road vehicles. Robotic mapping and navigation are mature research areas (Thrun, 2005) and open-source software stacks are now available to enable a car to pick and follow a route around static environment (Kato et al. 2015). In contrast, the human factors involved in real-world driving in environments containing other road users are not well understood. Other road users include human drivers and pedestrians, and eventually other autonomous vehicles, which may or may not be able to communicate with one another. A recent study operated an autonomous minibus continuously on a real commuter route around La Rochelle, France and Trikala, Greece, for several weeks (Madigan et al., in preparation) and suggested that once the local population has learned that the vehicle was programmed to be perfectly safe and stop at any obstacle, pedestrians and other vehicles would regularly take advantage of the AV to push in front of it. In some cases pedestrians did this simply for fun – delaying the vehicle and its real-world commuter passengers on their way to work. Human drivers do not simply give way to all such threats. Instead they appear to use their knowledge of and predictions about the psychology of other road users to interactively negotiate for priority. This can include dominant behavior – driving straight toward a pedestrian at full speed, or even accelerating, to encourage them to yield and return to the pavement; or polite behavior – backing off to allow them space to cross. A well-known scenario is for two pedestrians to meet similarly and make several attempts dodging from side to side in order to agree on how to pass. Two polite road users can be almost as dangerous as two dominant ones – by continually trying to yield to one another they get closer to a collision without the decision being made.

A recent study (Fox et al., 2018) proposed an abstract mathematical model of two agents – which may be AVs, human drivers, and pedestrians - - meeting at an unsigned intersection and negotiating for priority. This model is deliberately simple and intended as a foundation whose structure can be extended with many additional details. In addition, it contains free parameters decrypting human preferences. The present study illustrates how such parameters can be fit to human data as a method of measuring behavior which combines empirical Game Theory and a Bayesian Gaussian Process analysis.

Game theory of human driven vehicles is used extensively in macroscopic traffic modeling via Wardrop equilibrium in flow networks (Bolland et al., 1979) with focus on route selection in large, economy-like, markets of many road users rather than microscopic pairwise interactions. Where game theory has been applied to

Proceedings of Measuring Behavior 2018, (Manchester, UK, 6-8 June 2018).
Editors: Spink A.J. *et al.* www.measuringbehavior.org

pairwise traffic decisions, it has mostly been at the level of simple single-shot games as reviewed in (Elvik, 2014). In a few cases such as lane-changing (Meng et al., 2016; Kim and Langari, 2014) and merging (Kita, 1999) it has been extended to sequential games as used here, but not for AV-pedestrian interactions as here.

2. Methods

2.1 Sequential Chicken model

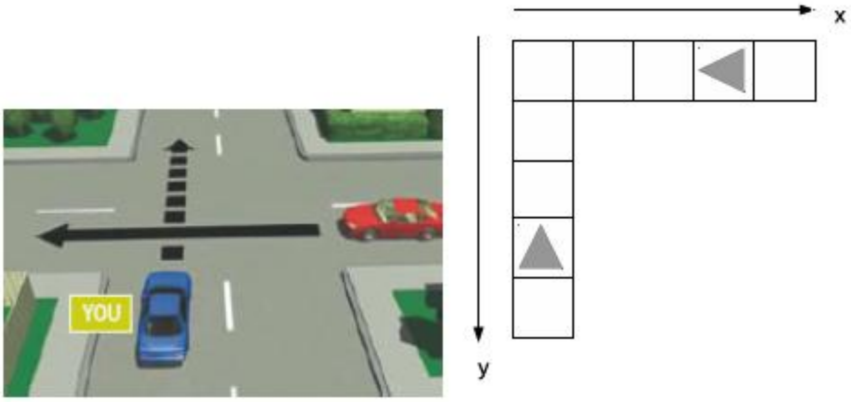


Figure 1 (left): two agents negotiating for priority at an intersection. (right): game theory model of the scenario.

The "Sequential Chicken" model of (Fox et al., 2018) represents a simplified version of the class of scenarios shown in fig. 1 (left) and fig. (2). In these scenarios, two agents (which may each be an AV, human driver of various types, and/or pedestrian) called *Y* and *X* approach the same intersection. If neither agent yields, they will collide and each receive a negative utility U_{crash}. Otherwise, they each reach their destination at some time delay, *T*, from the start of the game. The model assumes a linear value of time, so the total (negative) utility of an agent arriving at time *T* is $-TU_{time}$. The basic model assumes both agents share the same parameters U_{crash} and U_{time} and both know this to be the case. It assumes that no lateral motion is permitted, and that there is no communication between the agents other than seeing each other's positions. It discretizes positions into a plus maze structure of grid squares as in fig. 1 (right), and discretizes time into turns in which both players must select from two speeds, of either 1 square per turn or 2 squares per turn. The model is similar to a board game where each turn, both players make their choices in secret then and reveal them together before acting on them.

The model is called "Sequential Chicken" because its theoretical optimal strategies are solvable using Game Theory, and related to the famous game theory model of Chicken. In (non-sequential) Chicken, two drivers *Y* and *X* face each other and choose (as a single decision) whether to drive straight at each other or swerve away[1]. Swerving makes one the nominal "loser", but in practice, penalties for both are much worse if neither swerves and they collide. The game is considered a draw if both swerve. In game theory, this is represented by a payoff matrix, showing numerical utilities to the two players as in table 1.

$Y \setminus X$	a_X=*swerve*	a_X =*straight*
a_Y=*swerve*	(0,0)	(-1, +1)
a_Y =*straight*	(+1, -1)	(-100,-100)

Table 1: The game of Chicken.

[1] It is convenient to use *Y*, *X* ordering throughout so that matrix indices $M_{y,x}$ correspond to visualizations.

Standard Game Theory provides (Nash) *equilibria* for games in this matrix form, which are configurations of action selections such that neither player would change if they knew the other's choice. (Fox et al., 2018) provide arguments for selecting between these equilibria in real-time traffic situations so as to choose the most "rational" strategy. In general, the best strategies are probabilistic, specifying that each player should choose between the two actions with certain probabilities. (If they were deterministic, then each player could know the other's move in advance and exploit that knowledge, contradicting their supposed optimality.)

Chicken is a one-shot game – both players make just one decision then act on it. Sequential Chicken can be viewed as a sequence of one-shot (sub-)games, which can be solved similarly. Write (*y*,*x*,*t*) to describe the player's current locations as in fig. 1 (right) at turn *t* and let their available actions be the speed choices $a_Y, a_X \in \{1,2\}$. Rather than winning, losing, or crashing, the outcomes of the game at state (*y*,*x*,*t*) are that the players reach a new state, ($y+a_Y$, $x+a_X$, $t+1$). This may be a "game-over" state, where either one player has reached the intersection and/or a crash has occurred, or it may be another intermediate state leading to another game in the sequence. Define the *value* $v_{y,x,t} = ({v_{y,x,t}}^Y, {v_{y,x,t}}^X)$ of state (*y*,*x*,*t*) as the expected utility of the game to each of the two players, assuming that both play optimally from the state. Then the sub-game at time *t* can be written as a standard game theory matrix,

$$v_{y,x,t} = v(\begin{bmatrix} v(y-1,x-1,t-1) & v(y-1,x-2,t+1) \\ v(y-2,x-1,t+1) & v(y-2,x-2,t+1) \end{bmatrix})$$

which can be solved using recursion, Game Theory, and equilibrium selection to give values and optimal strategies at every state (given in Fox et al., 2018). Under a mild approximation (Fox et al., 2018) the solution may further drop all dependency on time *t* so that state values and optimal strategies becomes functions only of the player positions (*y*,*x*).

The model shows that if the two players play optimally, there must exist a non-zero probability for a collision to occur. This corresponds to the intuition that a perfectly safe vehicle which yields in every case will make no progress. Players must include some credible threat of collision to avoid being taken advantage of.

Figure 2: A "Barnes dance" or "scramble" crossing between Hollywood and Highland, Los Angeles. Real-world pedestrian-pedestrian interactions similar to the experiment occur between users of the two diagonal crossings.

2.2 Experiments

The model allows for computation of optimal strategies given parameters (U_{crash}, U_{time}) but does not specify numerical values of these human preference parameters. Eventually we would like to measure such parameters from real-world traffic data such as CCTV of vehicles and pedestrians on roads. The present study is intended to illustrate only the method for measuring such behaviors rather than report results, so instead uses a simplified proxy. Rather than interact physically, pairs of human subjects play a board-game version of Sequential Chicken as above. Their moves are recorded and used to show how to fit the parameters.

Sixteen participants aged from 22 to 48 were divided into 8 groups of 2 players. Each group's experiments took place with players and an experimenter seated around a table in university room and last for about 10 minutes. All experiments were conducted in accordance with University of Lincoln research ethics policy.

Natural game. Two players, one designated as player Y and the other as player X were sitting around a table, they were asked to play the chicken game on a plus-maze shaped board with discrete space as shown in Figure 1(right). Player Y was starting from y=10 and player X from x=10 such that they were both starting 10 squares before the intersection. Players were first told that their goals were to avoid a collision and to be the first to cross over this intersection as quickly as possible as if they were trying to commute to their office to work there in a morning. Players were each given two cards containing the numbers 1 and 2, and told that at each turn they should select one in secret then both reveal their cards and make their moves on the board simultaneously.

Chocolate game. After playing 3 natural games, each group played a further 3 games in which specific rewards were specified in advance. Players were told that the winner is the first to pass the intersection, and will receive two chocolates (Celebrations, Mars Inc.), while the second to pass the intersection will receive one chocolate, that both will receive no chocolates if there were any collision. Chocolates were provided after each game as in the instructions.

2.3 Measuring behavior

We wish to infer the parameters $\theta = (U_{crash}, U_{time})$ from the observed data, D, assuming that players play Bayes-optimally as prescribed by the model, M. By Bayes' theorem, the posterior belief over these is given by,

$$\mathrm{P}(\theta|M,D) = \frac{\mathrm{P}(D \mid \theta, M)\mathrm{P}(\theta| M)}{\sum \mathrm{P}(D| \theta',\ M)\mathrm{P}(\theta'| M)}$$

Given a set of possible θ values. assume a flat prior over θ, so that,

$$P(\theta|M,D) \propto P(D \mid \theta, M).$$

Hence we can work only with the data likelihood term, $\mathrm{P}(D \mid \theta, M)$, which is given by,

$$\mathrm{P}(D \mid \theta, M) = \prod_{game} \prod_{turn} \mathrm{P}\big(d_Y^{game,turn} \,|y, x, \theta, M\big)\, \mathrm{P}\big(d_X^{game,turn} \,|\, y, x, \theta, M\big)\,,$$

where $d_{player}^{game,turn}$ are the observed action choices, y and x are the observed player locations at each turn of each game. The action probabilities $P\big(d_Y^{game,turn} \mid y, x, \theta, M\big)$ can be read directly from the optimal strategies.

The above assumes that model M is perfect. However, M assigns zero or near-zero probabilities to many events which are found to occur in the data. Uncontrolled, this can be catastrophic for any Bayesian model of real data. We know that real humans may behave near-optimally but can make mistakes in perception (of the state of the game, such as the values of y and x) and in action selection. Thus, we weaken the model used in the analysis to a noisy but more forgiving version M'. This is a mixture model which generates each observation according to M with some probability s, and generates unbiased random actions otherwise,

$$P(D \mid \theta, M') = \prod_{game} \prod_{turn} [(1-s)\mathrm{P}\big(d_Y^{game,turn} \,|y, x, \theta, M\big)\, \mathrm{P}\big(d_X^{game,turn} \,|\, y, x, \theta, M\big) + s(\frac{1}{2})].$$

These equations are unlikely to have analytic solutions for θ, as they include elements of the optimal strategies computed via a long game theoretic process. To compute with them, we may sample various values of θ and observe $P(D \mid \theta, M')$. In practice, we work in log-space to reduce numerical errors, and consider $\log P(D \mid \theta, M')$.

Given a sample of these log-likelihoods, we may then fit a Gaussian Process (Rasmussen & Williams, 2006) regression model to them. Assuming that log-likelihood is a smooth function of the parameters, this provides estimates of the log-likelihoods at locations between the sampled values and can also show the uncertainty in these estimates. A Radial Basis Function kernel was used in the Gaussian Process model, with variance set to a minimum after obtaining smooth likelihood surfaces.

3. Results

Before performing the full analysis, we show in table 1 a rough overview of the game results. More collisions have occurred in the chocolate game than in the natural game. Define "delay" as the distance in boxes, or time in seconds, between the arrival at the intersection of the first player to arrive and the second player to arrive. The average delay in distance (box) is about 1.368 for the natural game but is reduced in the chocolate game (1.125). The time delay is computed following the formula: $\boldsymbol{d} = \boldsymbol{vt}$ where $\boldsymbol{d}$ is the distance, $\boldsymbol{v}$ the speed is assumed to be the maximum (2) and $\boldsymbol{t}$ is the time delay.

Type of Game	Number of Games	Collisions	Average Delay (box)	Average Time Delay (s)
Natural	24	5	$\frac{26}{19} \approx 1.368$	$\frac{26}{19*2} \approx 0.684$
Chocolate	24	8	$\frac{18}{16} = 1.125$	$\frac{18}{16*2} = 0.5625$
Total	48	13	$\frac{44}{35} \approx 1.257$	$\frac{44}{35*2} \approx 0.628$

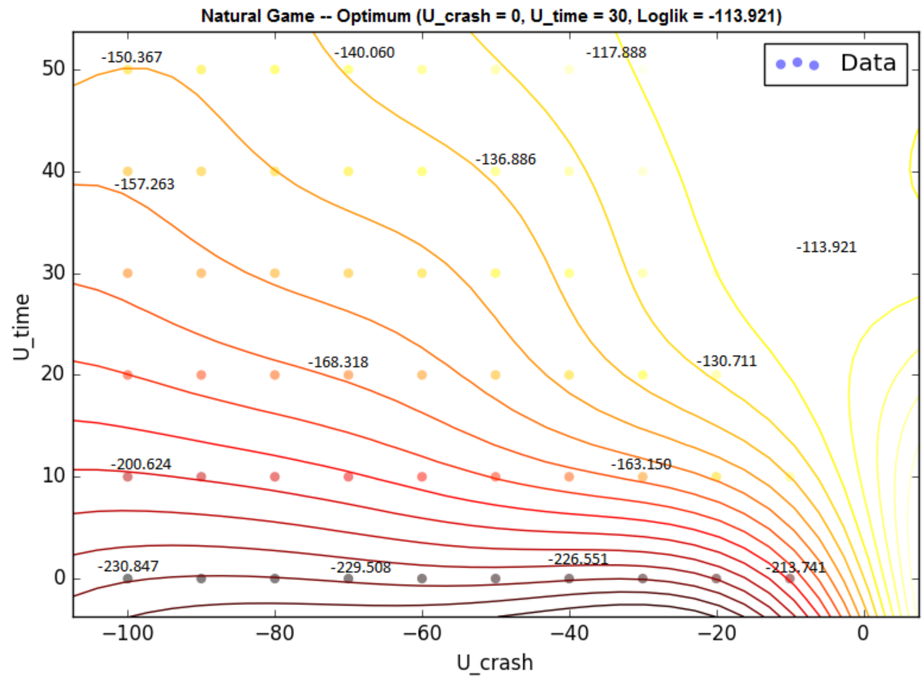
Table 2: Number of collisions and average time delay per game

Fig. 3 shows the results of fitting the Gaussian processes to likelihoods sampled along a regular grid, using s=1/10. The surfaces have a roughly radial structure. By visual inspection, the peak of the smoothed Gaussian Process surface in the natural case appears to lie somewhere in the region of $(U_{crash}, U_{time}) = (0,30)$ and for the chocolate case, around (0, 30). Fig. 4 shows a 1D slice through the 2D surface of Fig. 3(natural) which provides scale and uncertainty information.

4. Discussion

We have illustrated a new method for measuring and modelling human road user behavior, which is intended as part of a method to imitate or exploit such behavior in autonomous vehicle controllers. The model can be trained on human behavior then transferred into a controller to specify new robotic behaviors.

The particular experiment here has inferred posterior beliefs about the parameters of crash utility and time delay for human participants, in a simplified and artificial version of a real-world intersection negotiation task. We do not consider either the model or the experimental setup to be especially realistic versus real-world interactions, but they are sufficiently complex to illustrate the general method. This method is presented as an invitation for future work to extend both the model (e.g. with continuous positions and speeds, uncertain beliefs in opposite player utility function, lateral motion) and the experimental data collection (e.g. use of "big" real-world CCTV video of road-user interactions) stages to be more realistic.



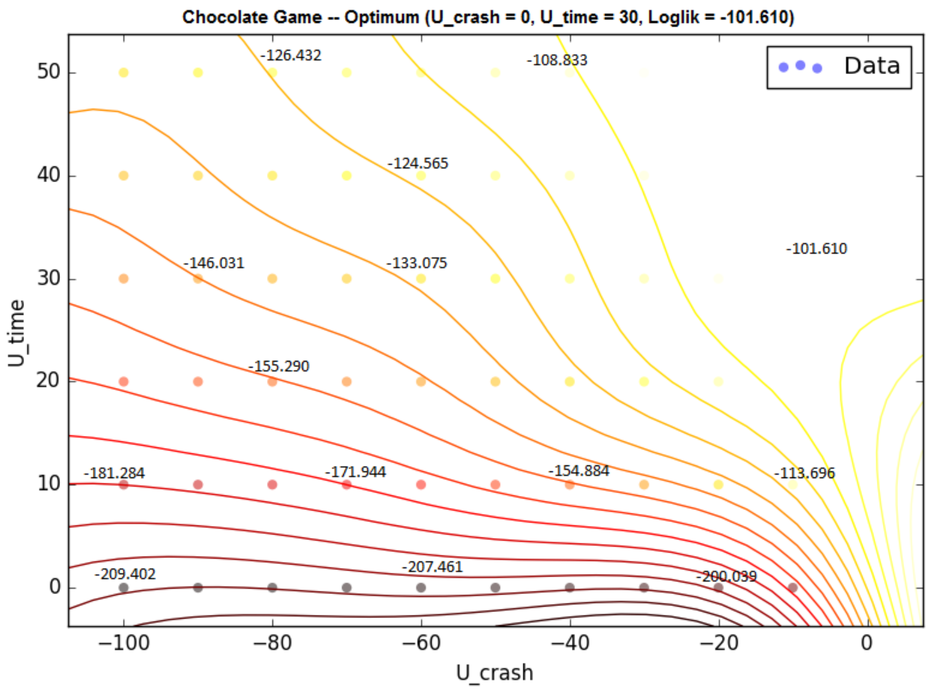


Figure 3: Gaussian process regression shows the likelihood (and hence, posterior belief) surfaces for the parameter values(U_{crash}, U_{time}), for the natural and chocolate experiments. The sampled parameter locations are shown by the dots. (This figure should be viewed in color).

Some conclusions from the results of the particular experiment analyzed which may be of interest, despite its artificial setting: From the viewpoint of real-world interactions, it is surprising to see a preference towards time-saving over crash avoidance in the players. In the real world, the cost of crash is usually millions of times more than the value of a second of commuting time, but here the players behave the opposite way. The effect is more pronounced in the chocolate game than in the natural one. This may be because the board-game-like nature of

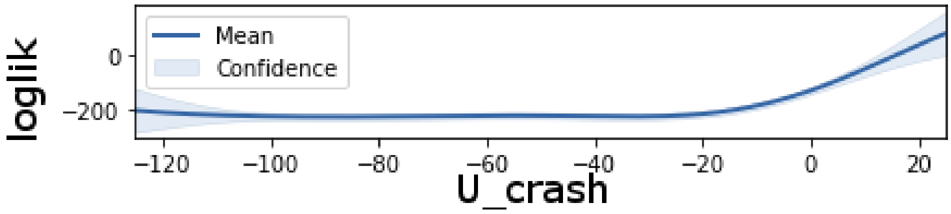


Fig. 4. Horizontal slice though the Gaussian Process of fig. 3(natural), at U_time=10, showing mean and standard deviation of belief about the log likelihood, $\log P(D \mid \theta, M')$ over the parameter space.

the experimental setting encourages the players to think in terms of winning and losing rather than just optimizing their personal utility. In the chocolate game this is made to be explicit by rewards the players specifically as winner and loser in the interaction. This leads to more collisions in this more competitive game.

The shapes of the likelihood surfaces are roughly radial, which suggest that it may only be the ratio between the two parameters which is important rather than their absolute values. Possibly the angles of the contour lines define equivalence classes of parameters. The number of experiments performed was quite small, and future work including further experiments may produce more peaky posterior distributions.

By using this method to fit more realistic models and data to human pedestrian-vehicle and vehicle-vehicle interactions, we can understand and predict behavior of road users to allow autonomous vehicles to more optimally interact with them. The present model requires a non-zero probability of collision to exist as a threat in all interactions. Future work could extend the model with additional, smaller negative utilities such as humiliating annoying pedestrians with horns or spraying them with water jets, in order to reduce or eliminate the necessity of actually hitting them whilst still allowing autonomous vehicles to make progress.

Acknowledgements

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreements No 723395 (InterACT) and No 645376 (FLOBOT).

References

Bolland, J., Hall, M., Van Vliet, D., and Willumsen, L. (1979). Saturn: Simulation and assignment of traffic in urban road networks. In Proceedings of the International Symposium on Traffic Control Systems, Berkeley, Calif. D, volume 2, pages 99–114.

C.W. Fox, F. Camara, G. Markkula, R.A. Romano, R. Madigan and N. Merat. (2018) When should the chicken cross the road?:Game theory for autonomous vehicle - human interactions. In Proceedings of VEHITS 2018, Funchal, Madeira, Portugal.

Elvik, R. (2014). A review of game-theoretic models of road user behaviour. Accident Analysis & Prevention, 62:388–396.

Kato, S., Takeuchi, E., Ishiguro, Y., Ninomiya, Y., Takeda, K., and Hamada, T. (2015). An open approach to autonomous vehicles. IEEE Micro, 35(6):60–68.

Kim, C. and Langari, R. (2014). Game theory based autonomous vehicles operation. International Journal of Vehicle Design, 65(4):360–383.

Meng, F., Su J., Liu, C., and Chen, W.-H. (2016). Dynamic decision making in lane change: Game theory with receding horizon. In Conference: 11th UKACC International Conference on Control.

Rasmussen, C. E., & Williams, C. K. (2006). Gaussian processes for machine learning. 2006. *The MIT Press, Cambridge, MA, USA*, *38*, 715-719.

Thrun, S., Burgard, W., and Fox, D. (2005). Probabilistic robotics. MIT press.

Madigan, R., Nordhoff, S., Fox, C., Ezzati Amini, R., Louw, T., Wilbrink, M., Schieben, A., and Merat, N. (in preparation). Understanding interactions between automated vehicles and other road users. Manuscript in preparation.