Game Theory Models for Pursuit Evasion Games

Mohammad Emtiyaz Khan
Department of Computer Science
University of British Columbia
emtiyaz@cs.ubc.ca

Abstract

In a pursuit evasion game, the pursuer tries to capture the evader while the evader tries to prevent this capture. A classical approach is to model this game as an infinite differential game. In this report, we model a pursuit-evasion game as a finite extensive form game, and show that the differential game is a generalization of this game. We prove that a pure-strategy Nash equilibrium always exists in an extensive game formulation, which is not always possible in differential games. Next we show that a finite-time termination issue can be resolved by modifying extensive form as a repeated game. Furthermore, we prove the existence of a Nash equilibrium in this repeated game. Finally, we discuss the relationship between repeated game and a popular approach known as probabilistic pursuit-evasion game.

1 Introduction

A pursuit-evasion game (PEG) consists of two players, a pursuer and an evader. The pursuer tries to *capture* the evader in some sense, while the evader tries to prevent this capture. A PEG presents a mathematical abstraction of many practical problems, e.g. surveillance using mobile robots where a swarm of robots act as a pursuer trying to capture the evader, or a guided missile chasing an aircraft.

An archetypal example of a PEG is known as the Homicidal Chauffeur game [1]. In this game, the driver of a car attempts to knock down a pedestrian, who, of course, does not wish to be flattened. The car can move faster than the pedestrian, but the pedestrian can maneuver itself better than the car. The question usually asked is “what is the best strategy for the pursuer (the car) and the evader (the pedestrian) to follow in order for each to achieve their conflicting goals?” There are many other versions too, for example the “Lion and Man problem”, in which the lion wants to eat the man and the man wants to save himself, or the “Obstacle tag game”, where one player tries to tag the other player.

Previous Approaches Because of its relevance to many real-world problems, PEGs have been extensively studied. As the problem fits well into the control theoretic framework, the usual approach has been to use algorithms from control theory. The most popular approach is based on differential games, wherein a differential motion model is assumed for both pursuer and evader [1] [2]. A saddle-point equilibrium is found by solving Isaac equations, which is equivalent to solve for min-max value. The analysis is quite involved as the game has continuous time and space variables. However, it's assumed that both players know about their own position as well as each other's position. It is stated that the differential game is an infinite perfect-information zero-sum game. A detailed description is given in Section 3.

The assumption of *perfect information* is relaxed in a more recent approach called probabilistic pursuit evasion games [5] [4]. In this approach, game is discretized in space and time. In addition to this, both the pursuer and the evader are assumed to have an uncertain measurement of each other's position and perfect measurement of their own positions. The game is modeled as an imperfect information Markov game and it is shown that a *one-step Nash equilibrium* exists for this model. A brief review of this method can be found in Section 4. There is another approach based on worst case analysis [7], but we will not cover these approaches as they are not directly related to game theory.

Contributions We can see that the formulations in previous approaches are closely related to game theory. However all the previous approaches have been motivated by a control theory perspective. In this report, we take a game theory approach to model the game. Our main focus is to model the PEG using simple game theory models, and unify the results obtained in the previous approaches. We also discuss the existence of solution concepts using well-known results from game theory. We do not wish to give methods to find solution concepts, however we will comment on a few possible approaches, wherever possible.

The report is organized as follows: In Section 2, we present an extensive form game model of PEG discretized in time and space, and show the existence of a Nash equilibrium. In Section 3 we show that the differential games are a generalization of extensive form games. We also discuss two important issues with differential games and show that the extensive form games provide important insights into resolving these issues. In Section 4 we show that one of the issues can be addressed by modeling PEG as a repeated game. After a brief discussion of the relationship between repeated game and probabilistic PEG, we conclude in Section 5.

2 PEG as an Extensive Form Game

In this section, we model PEG as an extensive form game. For simplicity, we quantize the game in both space and time. The field of action is assumed to be a bounded two-dimensional plane with discrete cells (Fig.1(a)). At every time-step, each player can move from one cell to another cell. This constitutes an action. An example of the possible action set is shown in Fig.1(b), where the pursuer can move left, right, up, down or stay where it is. Similarly, the evader can move to eight consecutive positions or stay at the current cell. If we code these actions as a set $\{0, 1, \ldots\}$, then the game can be written as an extensive form game (Fig.1(c)). We now formally define the game.

Notations Let us say that the field of action is of size $L \times W$, and X denotes the set of all discrete cells: $X = \{1, \ldots, W, W + 1, \ldots, LW\}$. Let t denote

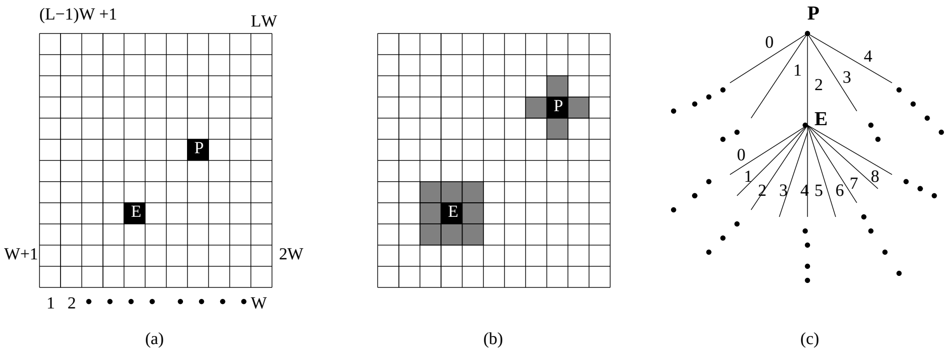

Figure 1: (a) A rectangular grid of size $L \times W$ with cells (b) An example of action space, in which the pursuer can move to four positions or stay wherever she is. Similarly the evader can move to 8 positions or stay at the current cell (c) Extensive form game, with action space coded as set $A_p = \{0, 1, \ldots, 4\}$ and $A_e = \{0, 1, \ldots, 8\}$.

the discrete-time which can take non-negative integer values. The set $N = \{p, e\}$ denote the set of players. The positions of both players at time t is denoted by $x_t = (x_t^p, x_t^e)$, where $x_t^p, x_t^e \in X$. Initial position x_0 is known to both the players. We also assume that both pursuer and evader have perfect knowledge of their own positions as well as each other's position. This assumption is just for simplicity and will be relaxed later on in Section 4.

Action Space The pursuer can take an action from an action set defined as $A_p = \{a_1^p, \ldots, a_n^p\}$. For example in Fig.1(b) the set is $\{0, 1, 2, 3, 4\}$. Similarly the evader's actions set is denoted by $A_e = \{a_1^e, \ldots, a_m^e\}$. Given the current position and the action, the next position is given by a motion (transition) model. To understand it clearly, consider the example given in Fig.1(b). If A_p is coded as $A_p = \{0, 1, -1, W, -W\}$, then the next position will be given by $x_{t+1}^p = x_t + a_i^p$, which is a simple linear model. In general, motion models are the functions which map the current position and action to the position at the next time instant:

$$x_{t+1}^p = f^p(x_t^p, a_i^p) \qquad x_{t+1}^e = f^e(x_t^e, a_j^e) \tag{1}$$

where $f^p : X \times A_p$ and $f^e : X \times A_p$ are motion models. These equations can be combined into one equation and written more compactly as follows:

$$x_{t+1} = f(x_t, a_i^p, a_i^e) \tag{2}$$

where f combines f^p and f^e. Given the initial positions and the actions, a node of the extensive form tree is characterized by a sequence $\{x_0, x_1, \ldots, x_t\}$, and hence equivalently by a sequence $\{x_0, a_{i_1}^p, a_{j_1}^e, a_{i_2}^p, a_{j_2}^e, \ldots, a_{i_t}^p, a_{j_t}^e\}$ (i_t and j_t are actions taken at time t). A mixed-strategy profile can also be defined with a probability distribution over these strategies, but to keep the analysis simple we only consider pure-strategies.

Terminal nodes For terminal nodes we need to define the capture condition. Intuitively, the evader is said to be captured when the pursuer and the evader are close enough, i.e. when the distance between the pursuer and the evader is less than some threshold. Formally, given $d_{min} > 0$, game ends when $d(x_t^p, x_t^e) < d_{min}$, where d is a distance measure. For example, in our discrete example d can be the shortest path between the cells. A problem with this formulation is that it allows some sequence of x_t which will be infinitely long. For example if the pursuer and evader are far away initially and don't move at all, then the game will never end. To avoid such cases, we impose a finite time restriction. Hence the game ends when

$$\{d(x_t^p, x_t^e) < d_{min}\} \text{ or } \{t = T\} \tag{3}$$

where T is a known finite time when the game ends. This restriction holds for some games where both players have finite energies. However, it is too strong an assumption and we will relax this later on while modeling the game as a repeated game.

Utilities Finally we define the utilities for the terminal nodes. Note that the pursuer and evader have conflicting interests and it's a zero-sum game, so $u_e(Z) = -u_p(Z)$, where Z is a terminal node. One way is to assign a utility 1 to every terminal node at which the pursuer has captured the evader, and assign all the other nodes a utility of -1. This definition will treat all branches with capture equivalently. In most cases, however, terminal node with a small capture time (or equivalently smaller length) is preferred to a longer one. Hence a better utility will be the sum of a function of the positions at all the time steps:

$$u_p(Z) = \sum_{\forall t} g(x_t^p, x_t^e) \tag{4}$$

Note that this function should be strictly decreasing with time, only then the minimum capture time will maximize the utility. For example, g can be (negative of) a distance measure at each time instant. In this case a branch, which minimizes the sum of the distance between the pursuer and evader, will be most preferred. Note that Eq.(4) can be re-written as follows:

$$u_p(Z) = \sum_{\forall t} h(x_t, a_t^p, a_t^e) \tag{5}$$

We will use the following equation because this is the way utilities are defined in standard game theory.

Final Extensive Form Game This gives us a perfect-information extensive form game[1] which we summarize now (we use the same notation as [6], Chapter 5):

1. Set of players $N = \{p, e\}$.
2. Action profile $A = (A_p, A_e)$.
3. Non-terminal nodes H characterized by sequence $\{x_0, a_{i_1}^p, a_{j_1}^e, \ldots, a_{i_t}^p, a_{j_t}^e\}$.
4. Terminal nodes Z with game over condition given by Eq.(3)

[1] It's easy to see that the popular missing and delayed measurements case can be modeled with an imperfect information game, which will be a slight modification of this game.

5. Player function ρ, actions function χ and successor function σ are trivial.

6. $u = (u_p, u_e)$ as defined by Eq.(5)

Equilibrium A few interesting remarks can be made about the equilibrium of the above game. The following game is a finite, perfect-information, zero-sum extensive form game. Using Theorem 5.1.3 from [6] (Chapter 5), we conclude that a pure strategy Nash equilibrium exists for this game. This is a direct consequence of modeling the game as an extensive form game. We will see in the next section that in case of infinite games existence of a pure-strategy Nash equilibrium is not ensured always. It seems that a subgame perfect equilibrium can also be shown to exist and various algorithms like backward-induction, alpha-beta pruning and minimax [6] can be used. Application of sequence form is also a possibility.

3 PEG as a Differential Game

We now review the model presented in [1] which is called a differential game. A differential game uses the (differential) motion model of the players. Such motion model are of the following form:

$$\frac{dx_t}{dt} = f(x_t, a_p, a_e) \tag{6}$$

Here x_t is the *state* of both the players (contains position, velocity etc.), $t \in \mathcal{R}$ is a continuous time variable, a_p and a_e are actions from continuous action sets A_p and A_e. One can see that the model becomes quite involved when both space and time are assumed to be continuous. An intuitive understanding can be obtained by looking at the discrete time equivalent of the following model[2]:

$$x_{t+1} - x_t = f^{'}(a_p, a_e) \tag{7}$$

The above equation computes the next position as the sum of the previous position and a function of actions. Similarly the continuous case in Eq.(6) relates the rate of change of positions (difference of position in some sense, or speed) to the current position and actions. The next position can be computed by taking the integration on both sides of Eq.(6).

Differential Game A differential game problem is to determine the saddle point of the following function [2]: ,

$$J = \int_0^T L(x_t, a_p, a_e)dt + \phi(x_T) \tag{8}$$

where T is the finite-termination time, L and ϕ are some functions. It's easy to see the equivalence of the above function to the utilities defined in the extensive form game (Eq.(5)). The first term is just a continuous time version of the Eq.(5). An extra second term in Eq.(8) is due to the boundary conditions in the integral (and hence depends only on x_T). Hence we can see that the differential game is a generalization of the extensive form game for continuous time and

[2]For simplicity, here we assume that x_t contains only the positions

space. The only major difference is that the differential game is an infinite game.

Equilibrium It might seem surprising that despite being an infinite fame, differential games are more popular than any other model[3]. The reason behind the popularity of differential games is because of the availability of elegant solutions to find equilibrium. When the function L is a quadratic function, then the saddle point can be found by using theory of optimal control [2] [3]. Saddle point is equivalent to min-max value of the game [3] [6]. This problem can be simplified to Isaac equations [1], solution of which is well-known in control theory. We will not go any further into describing Isaac equations, and interested readers should consult the references. However we will now comment on two important issues of differential games.

Two important issues There are two issues regarding the differential game, as discussed in [3]. We now discuss these issues and compare it with the extensive form game formulation.

1. **Pure-strategy equilibrium may not exist** Differential game is an infinite zero-sum game. For the value of game to exist it must be true that $\max\min J = \min\max J$. However minmax theorem holds only for finite games (an informal example can be seen in [9]). As described in [3], in such cases pure-strategy equilibrium may or may not exist. In most of the literature it is assumed that the above condition is satisfied and a pure-strategy Nash equilibrium holds. In contrast to this, as shown in Section 2, a pure-strategy Nash equilibrium always exists for the finite extensive form game.

2. **Finite-time Termination** It is assumed that the game ends after a finite time. However in more general games it is not at all certain that the game will terminate. We will show in the next section that if we model the game as an infinitely repeated game, we need not make this assumption.

4 PEG as a Repeated Game

We now model PEG as a repeated game which is just a simple modification of the extensive form case. At each time instant players play a (reduced) two-step extensive form game as a stage game. The only difference is in defining utilities: Every stage game has utilities $h(x_t, a_t^p, a_t^e)$, as defined in Eq.(5). The question is how to define the overall reward?

Discounted Reward We can see that at every time instant, the game will end with some probability, and this probability is a function of time. Assuming that we know these probabilities, we can make use of folk theorem to say something strong about the equilibrium. If we denote the probabilities of game end by a sequence $\{\beta_t\}$, then the discounted reward can be written as follows:

$$\sum_{t=1}^{\infty} \beta_t u_t^p \tag{9}$$

[3]In fact, to the author's best knowledge, there is no work on finite game formulation of PEG

where u_t^p is the utility of the pursuer as defined earlier.

Equilibrium We now use folk theorem ([6], Chapter 6) to show that a Nash equilibrium exist and it is nothing but the Nash equilibrium of each stage game. First we know that in every stage game, the maxmin value will be equal to the minmax value (as it's a finite zero-sum game). Hence a stage game Nash equilibrium payoff is enforceable. It is trivial that it is feasible as well. Hence we can say that a payoff profile in the stage-game equilibrium is a profile in some Nash equilibrium of the infinitely repeated game with average reward. It is shown in [8], that under some conditions any discounted reward can be written equivalently as an average reward. Hence the earlier statement about Nash equilibrium will also hold for discounted reward. We conclude that a payoff profile from the stage-game equilibrium is a profile of some Nash equilibrium of the repeated game[4].

At this point, it is worth mentioning another approach where the assumption of perfect information is relaxed. The approach is called probabilistic pursuit-evasion games, and it is closely related to our repeated game formulation. Lots of research has been done from a control theory perspective [5]. However only the results in [4] are relevant from a game theory perspective.

Probabilistic PEG In [4] PEG is modeled as a partial information Markov game. This is very similar to the repeated game formulation, however there are a few differences because of Markov game. The deterministic transition function is now replaced with transition probabilities $p(x_{t+1}, x_t, a_p, a_e)$. The measurement is given by a observation probability function. For example if y_t^p is the observation of the pursuer, then $q_p(y_t^p, x_t)$ is the observation probability function. With this observation function, observation sets Y_p and Y_e available to each player are defined. Because the information available to both the players is not the same, the game becomes a non-zero sum game. It is shown that a one-step Nash equilibrium exists, which is similar to our repeated game formulation with Nash equilibrium in the stage game. It seems that our results can be extended for the probabilistic case to explain the results given in [4].

5 Conclusion

In this report, we presented models of PEG as an extensive form game and repeated game. We showed that the differential game is a generalization of extensive form game. We addressed two issues of existence of pure-strategy equilibrium and finite-time termination. We showed that these two issues can be easily dealt with using extensive form and repeated game formulation. Finally, we discussed the relation with probabilistic pursuit-evasion games.

Acknowledgment The author would like to thank Mr. Leigh Burbidge, Sauder School of Business, UBC, for his support in reviewing this manuscript.

[4]We haven't explicitly stated the conditions in [8], which need to be explored further, however it seems that the conditions will hold in most of the cases

References

[1] R. Isaacs. *Differential Games* Wiley, New York, NY, 1965.

[2] T. Basar and G.J. Olsder, *Dynamic non-cooperative game theory*, Academic Press, New York, 1982.

[3] Y. C. Ho, A. E. Bryson Jr. and S. Baron, *Differential Games and Optimal Pursuit-Evasion Strategies*, IEEE transaction of Automatic Control, Vol. AC-10 No. 4, 1965.

[4] Joao Hespanha, Maria Prandini, Shankar Sastry. *Probabilistic Pursuit-Evasion Games: A One-Step Nash Approach.* In Proc. of the 39th Conf. on Decision and Contour., volume 3, pages 2272-2277, Dec. 2000.

[5] Joao Hespanha, Hyoun Jin Kim, Shankar Sastry. *Multiple-Agent Probabilistic Pursuit-Evasion Games.* In Proc. of the 38th Conf. on Decision and Contr., volume 3, pages 2432-2437, Dec. 1999

[6] Y. Shoham and K. Leyton-Brown. *Multi Agent Systems* In press.

[7] L. Guibas, J.-C. Latombe, S. LaValle, D. Lin, and R. Motwani. *A visibility-based pursuit-evasion problem.* International Journal of Computational Geometry and Applications, 4(2):74-123, 1985.

[8] M. Hutter, *General Discounting versus Average Reward*, Technical Report, IDSIA, Available at "http://www.idsia.ch/idsiareport/IDSIA-11-06.pdf".

[9] Lecture Notes, "http://www.maths.lse.ac.uk/Courses/MA413/IIL1.pdf"