Representing differential games as combinatorial games

Thomas Forster

July 17, 2004

Preface

An important and embarrassing fact about the theory of games—not often remarked on—is that there are although we have the one word 'game' we have at least three mathematisations of the prescientific concept corresponding to it, and there appear to be no connections between them.

- There is a well-defined and well-understood study of what are called "combinatorial" games, which may be finite (chess, draughts, *etc.*) but may also be infinite, They have wide applications in logic and complexity theory, and are the chief focus of this meeting.
- Matrix Games (as in von Neumann and Morgenstern's classic text) arose in economics, but they now have wide applications in politics and biology.
- Finally, there are *differential games*, arising from control theory and dynamical systems, which include—*inter alia*—pursuit and evasion games of interest to designers of missile systems.

Given the skill and ease with which most humans can cope with quite hard problems concerning games (in contrast to the pig's ear they routinely make of comparatively straightforward exercises in probability such as the Monty Hall problem) one can tell that games are important and natural mathematical objects. It is therefore *prima facie* odd that we have no account of how these three kinds of mathematical objects objects can all be one kind of mathematical object, namely a game. In this note I will exhibit a possible way of dressing up differential games as combinatorial games of imperfect information.

Motivation

The theme of this meeting is combinatorial games (specifically, those combinatorial games that arise in logic and semantics) and it might seem at first blush that the endeavour to make differentiable games look like combinatorial games has little to say to this programme. However we should not forget that (one of the many) differences between natural languages and the formal languages we have invented over the last century or so is precisely that the well-formed formulae of natural languages are continuous (analogue) objects in contrast to the discrete objects that are the well-formed formulae of formalised languages. And it may well indeed turn out that the application—to natural languages—of the insights obtained in the study of semantics of formal languages might profit from the parallel endeavour of applying—to differential games—the methods we have developed for discrete (combinatorial) games.

Differential games

In a game of pursuit P (pursuer) is trying to catch E (evader). The situation is deemed to be described by "kinematic equations" (*"KE"*) with coefficients in *control space*, which has some parameters under instantaneous control by P and some by E. In plain language what is going on—in all the cases that interest us—is that the fate of P and E depends on the behaviour of a dynamical system described by differential equations. We will use the evocative jargon of control theory here: "Control" variables are those that can be "controlled" by some agent in which we are interested—in this case player P or E. "State" variables are not under anybody's direct control, but are needed to properly describe the space in which the dynamical system lives. Thus state space is the space of values of the evolving solution. (and we can properly speak of *"the"* solution since initial conditions will be specified for any play). The control space is the set of parameters that are under the control of one or other of the players. One can think of P and E sitting at a console each with a joystick over which (s)he has sole control. By reparametrising where appropriate we may think of the control space as constituted by some parameters under P's control and some under E's control, and all independent.

The characteristic feature of pursuit games is that when the dynamical system enters a certain set of states (which are those where P is within ϵ of E) then the game is over, for P has won.

Conventionally it is assumed that P and E can at any time reset their control variables to any values they choose. The analysis below holds even if the behaviour of the evolving system is not described by differential equations—as long the evolution of the system is uniquely determined by what P and E do to their joysticks. On the other hand it is hard to imagine any useful applications of an existence theorem for strategies for pursuit games where the paths are not at least piecewise integrable, so—as an aid to comprehension at least—one need think only of games where the initial conditions and the form of the differential equations are such that however and whenever P and E change the coefficients over which they have control, the trajectory they produce is piecewise integrable, that is to say, it consists of finitely many bits of solutions to differential equations of the same form, all concatenated.

The idea is any pursuit game G of this kind between P and E can in some sense be reduced to a combinatorial game G' between two players P' and E'.

We tend to think that differential games of pursuity have winning strategies. There are some pretty plausible existence theorems for winning strategies. (Buggers to prove tho'!) Why do we believe it? It may be that we are thinking of these games as objects of finite character. This means that somewhere at the back of our minds we have reparametrised the phase spaces and chosen control variables so that the only manner in which the players alter the settings of the control variables is discretely and asynchronously. A game that can be reparametrised in this way is a prime candidate for being thought of as an object of finite character.

This is a nontrivial step. But even once we have done it there still remains the problem of how to deal with the asynchronicity and turn the game into a genuinely alternating one.

An unwelcome subtlety arises from consideration of refractory periods for the two players. (For a brief period after it has fired, a nerve cell is unable to fire. This recovery time is known as the "refractory period".) Any realistic analysis of games of pursuit must take account of the fact that there are several distinct "reaction" parameters to be considered, for example:

1. The time taken to abort a parameter setting;
2. The refactory period after each setting;
3. The time taken to execute a reply to the opponent's move.

We are going to assume (1) is 0: players can abort at arbitrarily short notice. This restriction is not absolutely essential (the construction can procede without it) but it makes life simpler. The relation between (2) (the refactory period) and (3) (the reaction time) is important. If reaction time is longer than refactory period then by the time player A has decided what to do in reply to a parameter setting by the other player, another move has been made. Effectively this means that the game is not a game of perfect information, since by the time the reply is executed, things may have happened in the game of which the author of the reply was unaware when it was formulated. In general there is no reason to suppose that games of imperfect information are determinate, and if we are trying to prove a determinacy result, we will just have to assume that the reaction time is less than the refactory period. This of course diminishes the strength of the determinacy result which is to follow, but it is unavoidable: games which do not satify this condition are unlikely to be determinate.

Although this complication with refractory periods does not preclude reconceptualising a game of pursuit as a combinatorial game, it will mean that the combinatorial game that results is a game of imperfect information. I am endebted to Francisco Corella for this following illustration of how this imperfection-of-information can prevent us establishing the determinacy result we would like. In Corella's counterxample, P and E are both forced to choose after time t_0 whether to turn left or right. Things are cooked so that P catches E iff they turn the same way. Clearly it matters in this game who we take to be player `I`. Whoever is `I` loses, because they announce which way they will go and the other player chooses to go the other way but announces it too late for them to change. We will clearly have to think very carefully about simultaneous moves and abortions: the implications of this game merit further thought.

Another important but easily overlooked assumption we must make is that the settings of the control parameters by one player are events visible to the other. For similar reasons we will insist that the initial conditions are known to both players.

A *strategy* for such a game of pursuit is a function from points in state space to parameter settings. Since by hypothesis we can recover the current position in state space from the history of parameter settings we may regard strategies also as functions from history of parameter settings to a current (or forthcoming) parameter setting-with-a-date, and this is the point of view adopted here. It may seem obvious that such games should have strategies, but a difficult feature is the nondeterministic nature of the sequence of moves: the players do *not* necessarily alternate moves and, although they may make finitely many moves in succession, this does not make the game of pursuit in any sense a "starred" game of the sort well-known in discrete infinitary games, for the players can never know how many moves they will make at each stage.

The Auxilliary Game

The basic idea is to construct an auxilliary game in which the two players (`I` and `II`) do indeed make their moves in strict alternation. It will be a standard combinatorial game like chess or draughts. A move in the auxilliary game is not a parameter setting (as it were a movement of the joystick) but an announcement of intended future parameter settings. The infinite divisibility of time intervals is a complicating factor, but the basic idea that enables us to construct an auxilliary game in which the players move alternately is not hard to grasp. One can consider a toy version of our continuous-game problem which lacks the infinite-divisibility complication but exhibits that same disagreeable absence of strict alternation of the players' moves which obstructs the application of the recursions needed to apply the Gale-Stewart theorem.

We have a set $A \subseteq \mathbb{N}^\omega$ in mind. `I` starts writing down natural numbers, and procedes until he is interrupted by `II` (`II` has to accept the last natural number `I` produces at that move), who then writes down numbers until interrupted by `I`, and so on. The players thereby construct an ω-sequence s of natural numbers, and `I` wins $\longleftrightarrow s \in A$. Let us call this "pseudogame", G. What is a notion of winning strategy for it?

The obvious way way to simulate a pseudogame such as G with a proper game H with genuinely alternating moves is to devise H as follows: each player either announces precisely one integer or passes, subject to the proviso that no player may pass if the other player had passed on the immediately preceding move. Evidently we can think of a *discrete* pseudogame as a discrete game of this sort, but this technique is less suitable for simulating continuous (pseudo)-games—if only because time is not discrete—so (since they are the current object of study) we shall not examine it further: the technique we use on the toy version must be one we can use on the full-size version.

The proper notion of auxilliary game for this toy version is as follows: Player I announces an ω-sequence of natural numbers. Then player II accepts some non-empty proper initial segment of I's proposal, and replies with an ω-sequence of natural numbers of her own. I similarly accepts some non-empty proper initial segment of II's proposal, and replies with an ω-sequence of natural numbers of his choice. They procede like this, constructing a gradually growing string of "accepted" numbers, and eventually an ω-sequence, s, say. Then as before I wins $\longleftrightarrow s \in A$. This *is* a game in which the players move alternately, and clearly also there is an intimate relation between this game (let us call this H as well, and forget the previous H) and the pseudogame G of the last paragraph but one. We must now examine this intimate relation.

How do we turn strategies for H into strategies for G? Suppose II, say, has a strategy σ in H. What use is σ to II when she plays G? In order to use σ, II has to attribute moves in H to I, and pretend that she is playing a *partie* of H.

Initially, in any play of G, II has not attributed any move to I, though later on she has, in a way we shall see below. She lets I continue writing down natural numbers until he has written down a sequence s, of length n say, such that there is an end-extension s^* of s, of length ω such that—supposing this were a *partie* of H and that I had made the moves II has attributed to him so far—σ would tell II to accept the first n elements of s^* and retort with some other ω-sequence t^*. (We say "II truncates at n".) There must come a point at which I has written down such a sequence, for otherwise at the end of time I has written down an ω-sequence s^* to which σ does not provide a reply. But σ is total. So for some n, σ says "truncate s^* at n" and then in the *partie* of G that she is actually playing, II knows to interrupt I as soon as she has seen him play those first n numbers. There may well be continuum many such t^* and AC will be needed to choose one. When such a point is reached, II picks a move and *attributes* it to I for future use. II interrupts I and starts writing down the members of t^* in order, until interrupted by I.

The time has now come to apply this to pursuit games. Let us fix a pursuit game of kind G. From each *partie* of G we can recover a wellordered string of dated parameter settings (we can ignore the state variables, since by hypothesis we can recover them by piecewise integration), which we can call a *log*. The hypotheses I have imposed on the theorem ensures that we deal only with games for which a log always contains enough information to discover who has won.

Out of force of habit I shall notate the two players as "I" and "II" as is common in discrete games, rather than "P" and "E". I is the player who moves first in a discrete game. It will make no difference to the argument whether we identify P with I and E with II or P with II and E with I. I will also respect the divergent traditions of speaking of a *partie* of a continuous game but a *play* of a discrete game.

We now define an auxiliary game H, played by I and II. We will use variables X_i, (resp. Y_i) for settings of I's (resp. II's) control variables.

I's ath move is a sequence

$$M_a = \langle\langle X_{(a,i)}; t_{(a,i)}\rangle : 0 \leq\ i \leq\ n_a \leq\ \omega\rangle$$

with a odd. II's bth move is a sequence

$$M_b = \langle\langle Y_{(b,i)}; s_{(b,j)}\rangle : 0 \leq\ j \leq\ n_b \leq\ \omega\rangle$$

with b even.

$\langle X_{(a,i)}; t_{(a,i)}\rangle$ and $\langle Y_{(b,j)}; s_{(b,j)}\rangle$ are *acts*, and their second components (*dates*) are elements of $\mathbf{R}^+$. Thus $X_{(a,i)}$ is the ith set of parameter settings planned by I (if a is odd, by II otherwise) to be done in the ath move. (Remember that a move is an announcement of a (possibly infinite) string of parameter settings). They are scheduled to be executed at time $t_{(a,i)}$. We require for player I

(i) $1 + t_{(a,i)} \leq\ t_{(a,i+1)}$;

(ii) $t_{(a,1)} \geq 1 + t_{(a-2,k)}$; where k is maximal such that $t_{(a-2,k)} \leq\ s_{(a-1,1)}$;

(iii) $t_{(a,1)} >\ s_{(a-1,1)}$;

(iv) $X_{(a,i)} \neq X_{(a,i+1)}$ for all i, a;

(v) $t_{(a,1)} \geq r + s_{(a-1,k)}$ where k is maximal so that $s_{(a-1,k)} \leq t_{(a,1)}$.

and for player II

(i) $1 + s_{(b,j)} \leq\ s_{(b,j+1)}$;

(ii) $s_{(b,1)} \geq 1 + s_{(b-2,k)}$ where k is maximal such that $s_{(b-2,k)} \leq\ t_{(b-1,1)}$;

(iii) $s_{(b,1)} >\ t_{(a-1,1)}$;

(iv) $Y_{(b,i)} \neq Y_{(b,i+1)}$ for all i, b;

(v) $s_{(b,1)} \geq r + t_{(b-1,k)}$ where k is maximal such that $t_{(b-1,k)} \leq s_{(b,1)}$.

Some comments:

The sequence of acts in the moves is the string of parameter settings that the player proposes to do if not interrupted. The idea is that in the auxilliary game the two players alternately announce what they are going to do between now and the end of time. As soon as the other player has made a move, however, any plans the one player has announced for subsequent periods are aborted, and (s)he is allowed another move. This ensures that in the auxilliary game the players do indeed move alternately. (i) and (ii) express the fact that the players have a refractory period after acting: the appearance of the numeral "1" in (i) reveals that our time-scale is normalised so the refractory period is of length 1! $t_{(a-2,k)}$ is the last act of I that was executed (see definition of *log* below) and he must wait 1 unit of time before proceeding, II *mutatis mutandis*. (iii) requires that you must decide once for all, at each stage, what you are proposing to do if your opponent does nothing, and that you cannot interrupt them before they have made their next move, even though you may truncate a proper terminal segment of it. (iv) is a piece of housekeeping that prevents us from spuriously attributing acts to the players. (v) ensures that a period r must elapse between the last move of one player and the first move of the other. We could actually discard this condition, since the bound r appears only in order to make the auxilliary game one of perfect information, and there will be times when a player might realistically be able to make a reply to an opponent's move *before* r units have elapsed: for example if the reply was a move the player had been proposing to do anyway.

A move M may be an empty sequence. In that case the preceding move is acted out in full, and M is the last move in that play of H (even though the *partie* it determines might go on for ever—for example if it is a Win for E).

We must now define a map π from plays of H onto logs of G. π is what a player in G will use to recover a winning strategy in G from one in H. To do this we must show that to each play p of H there corresponds a unique log L. The log is a selection of acts chosen from the moves in p. We build it by approximation. L_0 is the empty sequence. L_{n+1} is obtained from L_n by deleting from it all acts whose second coordinate is greater than $t_{n+1,1}$ (n even) or $s_{n+1,1}$ (n odd) and adding on the end all the acts in M_{n+1}. The idea is that once a player in the auxilliary game has announced certain proposed parameter settings (*acts*) for various times in the future, (s)he is committed to all of them that have time subscripts earlier than the time of the other player's next move, but not committed to any with later dates. $t_{0,1}$ will have to be ruled to be 0, and $t_{-1,1}$ to be -1 to start the recursion. Conditions (ii) for each player will ensure that any entry, although it may be added and removed finitely many times, will nevertheless either belong to or avoid a terminal segment of $\langle L_n : n \in \omega \rangle$, and we use the obvious definition for L_∞, or L for short. L is our obvious choice for $\pi(p)$. Obviously, what we want now is that the player who wins the partie corresponding to $\pi(p)$ wins p. This completes the definition of H.

If P wins the log (considered as a *partie* of G) then this becomes apparent at some L_n, and we may regard L_∞ as L_n in those cases.

We will be making use of the fact that this construction is continuous in the sense that to each position (initial segment of a play) in H we can associate an initial segment of a log. This function will also be called π. If l is a an L_n then $l|t$ will be the initial segment of l of acts executed before t.

Note that we here make use of the fact that the players are assumed to be able to abort *at arbitrarily short notice* plans to make any given setting of control parameters. Although this is a serious constraint it is arguably realistic. If one has the feeling that this is not realistic this probably shows that one also has the feeling that these control variables are actually the *state* variables of some antecedent control system, namely the player. The answer then is to reparametrise using the control variables of that system. The same thinking can be applied to the refractory period, but in either case this must be weighed against the need for the act of setting a control parameter to be publicly visible. This is not the place for a full discussion of these matters, since my purpose here is primarily to exhibit the construction, not debate its applicability.

However we identify P and E with I and II it will be the case that whoever is playing P in H will win, if at all, after finitely many moves. This makes H an open or closed game (this means that the set of plays that are wins for I is open (resp. closed) in some topology) and the Gale-Stewart theorem tells us that such games have winning strategies. Unfortunately, as we shall see, this game is reduced to a game of imperfect information, so we cannot straightforwardly use Gale-Stewart.

Use of the auxiliary game

How does a player in G (A, say, rather than P or E, to avoid question-begging) make use of a strategy σ that A has (remember we have somehow identified P and E with I and II) in H? In fact all we need for the theorem is to show that a *Winning* strategy for the auxilliary game gives rise to a *Winning* strategy in the original (pseudo-)game, but it would be nice to have a construction that worked for arbitrary games. However we will here talk through only the derivation for winning strategies, since we do not need the general case and it has serious complexities.

First A looks at the log l of the game so far, or to be precise, at the pair $\langle l, t\rangle$ of l and the current time. A must associate this pair with a position p in H that could have arisen by A playing according to σ. (p is "compatible with σ") We use the fact that σ is winning to give an easy inductive proof that the constructions below are possible. A finds such a p in one of two ways, depending on whether A or B was the author of the last entry in the log. (There always is a last entry, because of the refractory period)

(i) If B was the author of the last entry, then A looks at any position p in H such that $\pi(p)|t = l$ and p is compatible with σ. Since σ is winning, by induction hypothesis there will always be such a p. A then plays $\sigma(p)$ in H. If the date of the first component X of $\sigma(p)$ is a lot later than t, there is the possibility that B will have moved before X is executed, but this does not affect the construction, it merely means that in such cases σ is unhelpful, and in effect says "in this position, await developments".

(ii) If A was the author of the last entry then A looks at any position p in H such that $\pi(p)|t = l$, p is compatible with σ, and the last move X in p was $\sigma(p \setminus \{X\})$. As before there will always be such a p, since σ is winning. This last move X in p will have been one of A's so in a sense it is not A's turn, but moves in H are *sequences* of moves in G. A simply continues with the settings he blueprinted in X. To deal with the case where the last two acts were simultaneous we adopt a convention on the basis of who was the author of the penultimate act, and if that was B then (i) else (ii). There is an obvious recursive extension to the case where the last $2n$ acts came in simultaneous pairs.

But if the **first** *acts are simultaneous, what then?*

Now we invoke the Gale-Stewart theorem to infer that there is a winning strategy for H. It is evident that the auxilliary game we are looking at here is an open game, for if P wins at all, this has become apparent at some finite stage in the play of H.

Concluding Remarks

A good question is: is the solution time-invariant? That is to say, does P have to look at his watch in order to know what to do, or can he just look at their address in state space? *Prima facie* he does, in effect, look at his watch, for the strategy is a map from histories to moves. On the other hand, we can argue that if there is a winning strategy for player X then there is a winning time-invariant one. For if P has a winning strategy that tells him to do something different each time some point comes round, P should look at what it tells him to do the *last* time they come round (and there *will* be a last occasion, for P wins, if at all, in finitely many moves), and do that the *first* time. An analogous construction works for E: she makes at each occurrence of the state the move prescibed by the strategy at the *first* occurrence of the state. The fact that the construction works depends on the game being closed/open of course!

The presence of the refractory period is the other crucial restriction that enables us to construct a discrete auxiliary game to which we can apply the Gale-Stewart theorem. If we weaken the requirement that each player have a refractory period after each setting, even to the condition that there be a finite interval between successive acts, then there arises in the auxiliary game (as currently defined) the possibility of plays of transfinite length—indeed of any countable length. It may still be possible to recover a log from a play of the auxiliary game, and even perhaps to integrate piecewise and find out who wins, but the difficulty will then be that games of transfinite length cannot be shown to be determinate in the way that games of length ω can.

The referee of an earlier version of this paper has pointed out to me that in fact earlier attempts have been made to obtain results in continuous games by applying the results of discrete game theory, notably by Mycielski [2]

also references therein). The most natural approach to continuous games is to approximate them by progressively finer finite games which tend to the desired continuous game in the limit, and this is what Mycielski does. A certain amount of work has to be done to show that the same player Wins the limit as Wins the approximations. Thus in this approach Mycielski imposes conditions of convexity and compactness which makes it quite distinct from the development here, which does not require any analysis: the discrete games constructed here are not approximations at all, they are closer to the *coverings* used by Tony Martin in [1]

References

[1] Martin D. A. [1975] A new proof of Borel Determinacy. Annals of Mathematics **102** pp 365-71.

[2] Mycielski, J. Theories of pursuit and evasion, Journal of Optimization Theory and its Applications **56**, pp 271-284).