An information theoretic formulation of game theory, II

Chris Goddard
November 27, 2018

Abstract

This short article follows an earlier document, wherein I indicated how the foundations of game theory could be reformulated within the lens of a more information theoretic and topological approach. Building on said work, herein I intend to generalise this to meta games, where one game (the meta-game) is built on top of a game, and then to meta-meta-games. Finally I indicate how one might take these ideas further, in terms of constructing frameworks to study policies, which relate to the solution of various algebraic invariants.

1 Foreword

Herein I intend to build on [8] in order to indicate how one can make sense of 'meta-games', such as simplified poker. Determination of an approach to computing strategies for optimal play is discussed. Finally I indicate a few potential consequences of these ideas, both to ethics, as well as to generation of new AI frameworks and AI safety, in terms of derivation of optimal policies (which will be the focus of a subsequent paper).

2 Preliminaries

2.1 Operators, Operator Algebras

Recall from [9] and [3] that a central idea to the study of information theoretic structures are statistical manifolds, which are characterised by signal functions f, g, that map from the space of metrics to a probability over same.

In particular, if say $f(m, a) = \delta(\sigma(m) - a)$, $g(m, a) = \delta(\tau(m) - b)$, then we can construct $\star(f; g) := fg$, and $\circ(f; g) := f \circ g$, to first order. Clearly these will also be signal functions. Now I claim that there is a way one can build up various abstractions from these over operator algebras, so that they correspond in a natural way to the prime number sequence as we ramp up the complexity.

- Consider for instance we are thinking about function spaces (second order). Then, to second order, we have operators $\star, \circ$, and $\star^{(2)}$, where $\star$ and $\circ$ remain the same, but $\wedge = \star^{(2)}$ operates as $\wedge(f;g) := f^g$, for 3.

- For function-function spaces, or third order, things start to get interesting, because, instead of merely having the idea of things operating on functions, we have the idea of operators acting on operators.

  In particular, we have operators $\star$, $\circ$, $\star^{(2)}$, and $\circ\star$ and $\circ^{(2)}$, for 5 operators in total, where $\circ^{(2)}$ is tetrated composition, and $\circ^g \star (f;)$ is defined as the operator that takes a signal function h, and determines $\star(f;h)$, but then we lift this to $\star^{(2)}(f;h), \cdots, \star^{(g)}(f;h)$, ie $\circ(\star;g)(f;h)$. Evidently if g is between 0 and 1 this notation is not quite correct, but the idea is that we lift $\star$ to g levels of subtlety.

- At fourth order, we have additional operators $\circ\circ$ and $\star\circ$, for 7 operators in total.

- At fifth order, we have operators $\star^{(n)}, n = 1,2,3$, $\circ^{(n)}, n = 1,2$, $\circ\star^{(n)}, n = 1,2$, $\star\circ^{(n)}, n = 1,2$, and $\circ\circ^{(n)}, n = 1,2$, for 11.

- At sixth order, we bump $\star^{(m)}\circ^{(n)}$, $n, m = 1,2$ for 13.

- At seventh order, we bump $\circ^{(n)}, n = 1,2,3$, $\star^{(n)}, n = 1,2,3,4$, and introduce $\star\star\star$, $\circ\star\star$, for 17.

- At eighth order, we introduce $\circ\circ\star$ and $\star\circ\star$, for 19.

- At ninth order, we introduce $\circ\circ\circ$, $\circ\circ\star$, $\star\circ\circ$, and $\circ\star\circ$, for 23.

- At tenth order, we bump $\circ^{(n)}\circ^{(m)}, n, m = 1,2$, $\circ^{(m)}\star^{(n)}$, $n = 1,2,3, m = 1,2$, $\circ^{(m)}\circ^{(n)}$, $n = 1,2,3, m = 1,2$, $\star^{(m)}\circ^{(n)}$, $n = 1,2,3, m = 1,2$, $\star\star\star^{(n)}$, $n = 1,2$ for 29.

- At eleventh order, we bump $\star^{(n)}$, $n = 1,2,\cdots,5$, $\circ^{(n)}, n = 1,2,3,4$ for 31 operators in total.

Note that this process is pseudo-inductive, in that, although one can construct the prime number sequence in this way, the complexity of the construction at stage n+1 depends on rules that include all rules for stages 1 through n.

But why is this important? Well, it turns out that the choice of levels of abstraction in a meta-game, and the dimensionality of a topological product that we might wish to pick, is intimately tied to the level of *subtlety*, that, is, *function-function-$\cdots$-function* space 'depth' that we are carrying out our analysis on. Here I construe depth or subtlety broadly as 'relating to the amalgamation and outworking of all inductive rules arising from emergent structure, and characterised by the prime numbers, up to said integer'.

2.2 Metagames - A couple of examples

We now turn to the idea of meta games, with potentially more than one game subsidiary to an individual meta game. For the prototypical example, we consider the following iterated game:

Example. (Transparent Random Draw). We have N players in our game. Each players starts with M chips. For a new round, each player puts in a chip. The dealer deals one of three hands to each player: *good*, *bad*, and *ugly*. These are laid face up on the table where everyone can see. Ugly beats Bad, Good beats Ugly, Good beats Bad, and all other options draw. The dealer tallies which coalitions defeated which other coalitions, and divides the winnings up accordingly. If there is no Ugly or Good on the board, players with Bad hands keep their chip; if not, the players with Ugly hands divide the chips, unless there is at least one player with a Good hand, in which case the players with Good hands divide the chips. If there are an odd number of chips, the house takes the remainder. Finally, if a player is reduced to zero chips, they are out of the game.

Transparent Random Draw is clearly a fairly simple iterated game. There is no strategy; it is essentially a play and hope kind of game. But we can consider the following generalisation:

Example. (Simplified Poker). Consider the game Transparent Random Draw as before, with the following additional rules:

- Each player may bluff. That is, they may choose to pretend to have a better hand than they actually do, they may pretend that they have a worse hand than they actually do, or they may choose not to reveal any information. We make the simplifying assumption that no other player can tell whether a player is lying or not about their position, nor whether they are telling the truth - i.e., they have no information as to why a player has taken an action other than that they did.

- There are four phases during a Round (Epoch). During each phase except the first, a transformation card will be revealed. The transformation card will rotate $0, 1, 2$ where 0 is bad, 1 is ugly, and 2 is good, by either 0, 1, or 2 spaces. So transform-0 will leave things unchanged, transform-1 will transform bad to ugly, ugly to good, and good to bad, etc. Transformations compose. Also, there are three transformations in total, and these all start unrevealed. Finally, there are 6 transformation cards in total - 2 of each type (the reason we limit this and specify multiplicity is so that players have some information on which to base strategy according to revealed information over the course of a Round).
- Each player has an additional three actions during a Round. They may raise by one chip, they may meet the current bid, or they may fold. The number of raises possible during a Round is limited only by the number of chips that the two wealthiest players still in the bidding hold.
- If there are no players left except for one, then they win the Round and no further phases take place. Otherwise, when the bidding has concluded, the players who are still in the bidding reveal their hands, and all transformations are revealed and applied. Chips are then assigned as per Transparent Random Draw, and the next Round begins.

Then this is what I will term an *iterated meta-game*. That is, there is a game underneath (random draw of good, bad and ugly hands) with a *second game built on top of it* (bluffing and raising). Evidently the objective of the game is the same as Transparent Random Draw - it is to walk away with all the other players' chips.

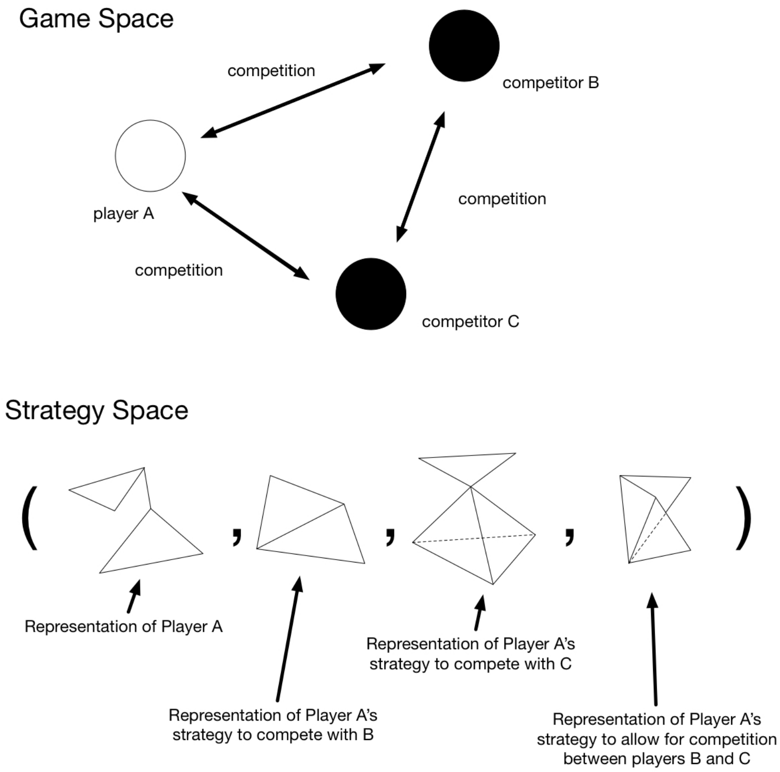
However, not everything is totally clearcut here. For instance, consider the following pair of examples:

Example. (Commuting). N players wish to travel from one part of a city to another. They have two modes of travel: bus and train. There is a graph of nodes with bus routes and train routes connecting the nodes. The objective is to travel from a given transport node A to an end transport node B as quickly as possible.

Example. (Commuting with Tickets). The same as commuting, but with a cost to travel between different nodes that differs from join to join. The objective here is to minimise cost of travel.

Then *Commuting with Tickets* is not a metagame of *Commuting*, as it is clear that a good strategy for the latter significantly interferes with a good strategy for the former. But how can we make this precise?

Example: A standard game



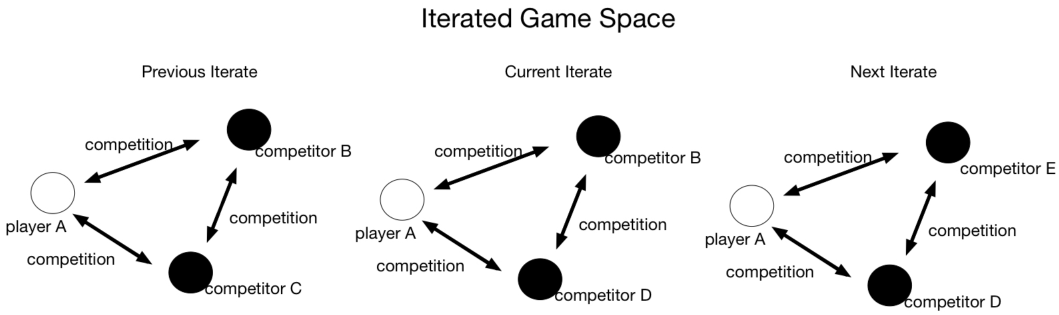
3 Decisions and Strategy in Meta Games

3.1 Formalism

So how can we faithfully represent the idea of a metagame? And, given that, how do we formulate an optimal strategy for a given metagame? We perhaps expect to have two nested information functionals / payoff functions. So, what might this look like?

Consider as before in [8], $h : \Delta^{(1)5} \to R$ as our iterated game. Here 5 copies are chosen for the following reason: consider as in the previous paper a decision for a coalition a, consisting of (a_0, a_1, a_2, a_3), where a_0 is a representation of the coalition in the game, a_1 is a decision given competition with coalition b, a_2 is a decision given competition with coalition c, and a_3 constitutes the shape of the decision for a adjusting for competition between b and c in the game.

For an iterated game, note that each coalition a must needs now *respond*, *act*, *Example: The structure of a standard iterated game*



and *anticipate*. Therefore, the decision vector is now a decision matrix, with $a_{-1,0}$ being the state of a in the previous iteration, $a_{0,0}$ being the state of a in the current iteration, and $a_{1,0}$ being the anticipated state of a in the next iteration. We define $a_{-1,i}, a_{0,i}, a_{1,i}$ for $i = 1, 2, 3$ in a natural way (being competition with b, c, and adjusted decision for competition between b and c in the previous, present, and future iterate respectively).

However, the coalitions competed with previously and in the future could actually be different. In particular, a will be competing with $b^{(-1)}$, $c^{(-1)}$ in the past, $b^{(0)}, c^{(0)}$ in the present, and $b^{(1)}, c^{(1)}$ in the future. But it turns out to be more natural to consider $b^{(-1)} = b^{(0)}$, and $c^{(0)} = c^{(1)}$, so that a competes with $b^{(-1)}, c^{(-1)}$ in the past, $b^{(-1)}, c^{(1)}$ in the present, and $b^{(1)}, c^{(1)}$ in the future. Why is this? Because information can only be carried over between iterations regarding competition, if at least one of the pairs remains the same.

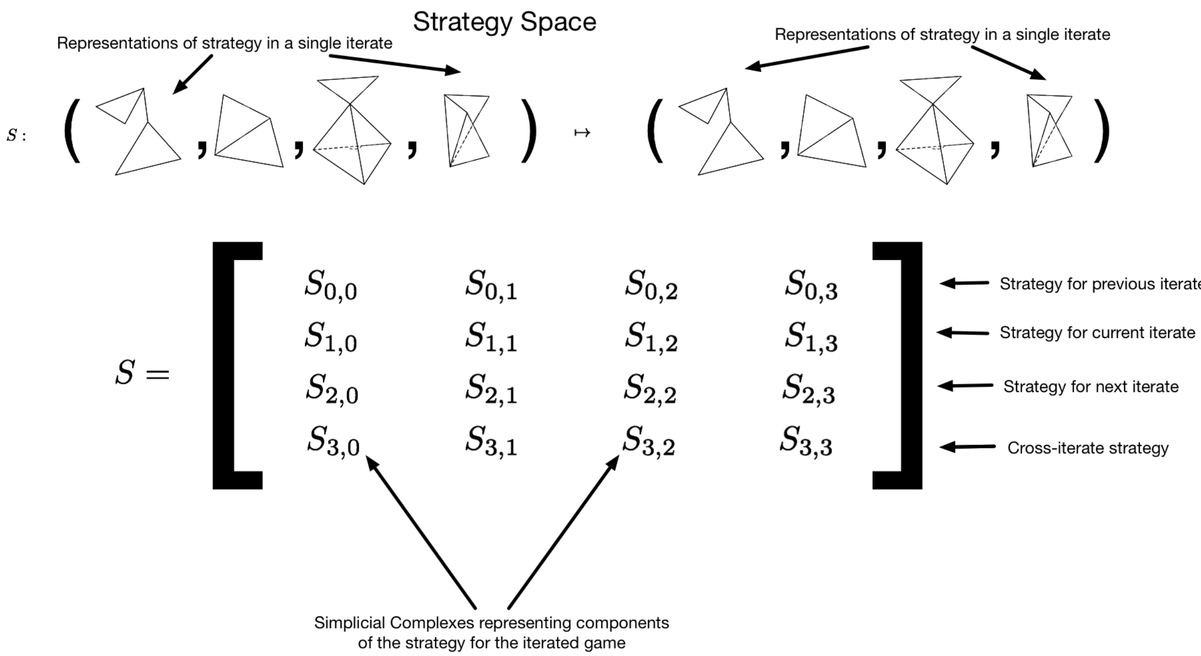
Consequently, generically we have a situation where we have 5 coalitions characterising our game over a decision matrix, or for elements of $\Delta^{(1)}$. Consequently the choice of five copies.

Now define $\mathcal{F}^{(0,1)} := \{h | h : \Delta^{(1)5} \to R\}$ as the set of all iterated games, and $\mathcal{F}^{(0,0)} := \{h | h : \Delta^{(0)} \times \Delta^{(0)} \times \Delta^{(0)} \to R\}$ as the set of all one-off games.

Then

$$k^{(0)} : \mathcal{F}^{(1,0)7} \to R$$

Example: Strategy within a standard iterated game



is a one-off metagame, where $\mathcal{F}^{(1,0)} := \{\phi | \phi : \mathcal{F}^{(0,0)} \times \mathcal{F}^{(0,0)} \to R\}$.

We justify the choice of 7, as intuitively, a one off metagame should be more complex that an iterated singleton game. To make this a bit clearer, we can consider the following: Let a be a coalition of players playing our metagame $\mathcal{G}$. Then $a := (a_{00}, a_{01}, \cdots, a_{03}, a_{10}, a_{11}, \cdots, a_{23}, a_{30}, a_{31}, a_{32}, a_{33})$ defines a 'point' taken by k. Here a_{00} represents the coalition for a. But what does a_{01} represent? a_{01} represents the strategy for a in the embedded game given the actions of coalition b, adjusting for the fact that these actions are within the confines of the simplified game. Similarly, a_{11} represents the strategy for a in the bridge / information linkage between the game and the metagame, given the actions of b. a_{21} represents the strategy for a in the metagame, given the actions of b. Finally, a_{31} represents the strategy for a, given the interaction of the adjudicator (who is operating outside the metagame itself) with the metagame, given the competitive actions of a with b. e.g., in simplified poker, this would be *the house*, and might be represented by *house rules* or *house rulings.*

a_{10} is the representation of a in the bridge, a_{00} is the representation of a in the game, a_{20} is the representation of a in the metagame, and a_{30} is the representation of a in the house ledgers.

Similarly for a coalition c competing with a, in such a way we define $a_{02}, a_{12}, a_{22}, a_{32}$. Finally we adjust for the competition of b with c by the decisions $a_{03}, a_{13}, a_{23}, a_{33}$.

So it is clearly as to why our coalitions a, b, and c should live in $\mathcal{F}^{(1,0)}$ as 'decision matrices', since they are defined over a function space of $\Delta^{(0)}$. But why should k take seven of them? The key to this observation is the idea of *the house*. The house may be running multiple metagames concurrently. Therefore, the way the house will interact with one game will generically be intimately tied to the outcomes of other metagames (say, if in simplified poker its pool of chips was running low).

Therefore, consider two metagames with competing triads a, b, c and a, e, f, with the dealer d wedged between them as a Bridge, and the house h representing the relationship with other games. Then, if we lift the idea of metagame from considering it as isolated instances between triads, to including the house, it becomes readily apparent that we must consider seven copies of $\Delta^{(1)}$ for things to make sense. In particular, the house h will have a decision matrix, with actions depending on how coalitions interact with the *house tables*.

Hence, proceeding to the idea now of an *iterated metagame*:

Consider $\mathcal{F}^{(1,1)} := \{\phi | \phi : \mathcal{F}^{(0,1)} \to \mathcal{F}^{(0,1)}\}$.

Then

$$k^{(1)} : \mathcal{F}^{(1,1)11} \to R$$

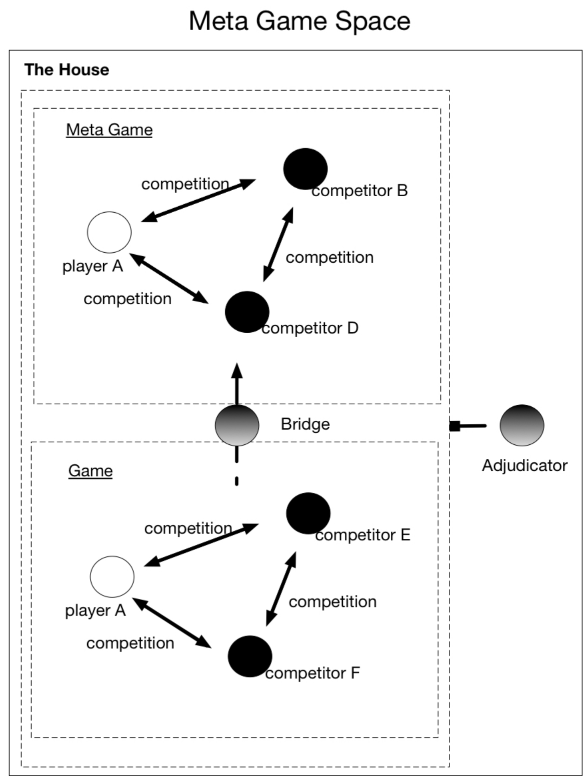
is an iterated metagame. We choose 11 copies because intuitively we expect an iterated metagame to be more complex than a one-off metagame. But to make a case for this, proceeding by analogy from before, note that each coalition a, just as before for an iterated game, must needs now *respond*, *act*, and *anticipate*. Response concerns the outcome of the previous iterations. Action concerns what to do this iteration. Finally, anticipation concerns what the coalition expects to happen in consequent iterations. So we expect our decision matrix to now be a decision tensor, with coordinates in past, present, and future.

Remark. Note that it makes sense for the next degree of abstraction to look two iterates into the past, and two into the future, as we have more coalitions involved!

So, while managed by h,

- a is competing with $b^{(-1)}, c^{(-1)}$ 1 iterate in the past (allowing for flow of information).

Example: The structure of a standard meta game

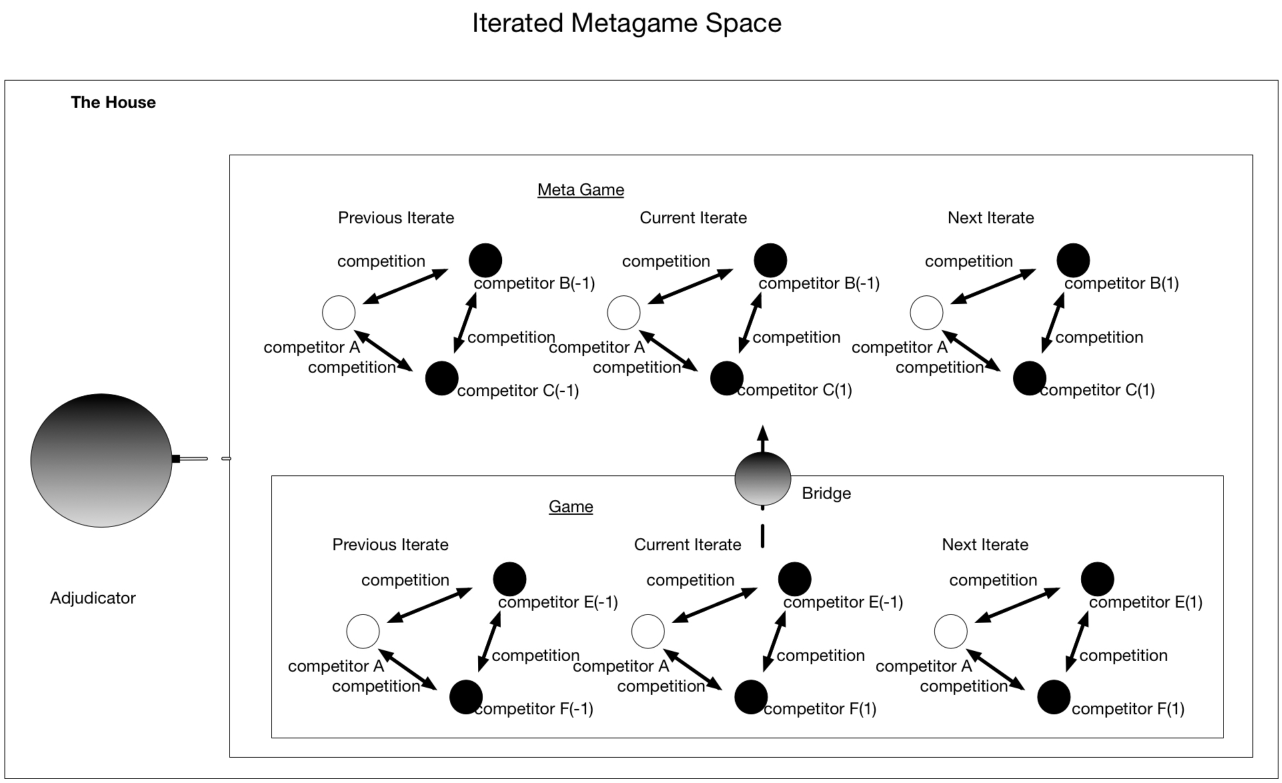


- a is competing with $b^{(-1)}, c^{(1)}$ in the present
- a is competing with $b^{(1)}, c^{(1)}$ 1 iterate in the future

(Note that this is a special case of a braid).

Therefore we have a decision tensor that has 4 more coalitions involved that for a one-off metagame (since this occurs for both copies of the metagame mediated by h), which makes for 11 coalitions required to specify the game. It is also clear that a, b, c belong in $\mathcal{F}^{(1,1)}$, since they are defined over a function space of the functions from $\Delta^{(1)}$ to itself (i.e., $\mathcal{F}^{(0,1)}$).

Example: The structure of a standard iterated meta game



Remark. (Emergent complexity matches the prime number sequence). Note that the analysis from game, to iterated game, to metagame, and then to iterated metagame, follows in essence the prime number sequence: 3, 5, 7, 11. This is no accident. In particular, note that in all of this we have the idea of one or more natural *pivots* in our structures. These correspond roughly to *signal functions* from the previous section, and, by analogy, our 'operators' are the analogues to the operators previously described. Hence we establish a natural connection between emergent complexity in the theory of games, and the prime number sequence.

In particular, if h is our information for an iterated game, then we compute an information for our iterated metagame ultimately as a function of this initial information; so that we have two levels of nested information structures, as in [7].

3.2 On structures in game theory

In an analogous way, we can proceed pseudo-inductively to build up natural structures over 'tangent bundles' for games, iterated games, metagames, and iterated metagames.

For games, we have a decision metric, $\sigma : T\Delta^{(0)} \times T\Delta^{(0)} \to R$, which determines how a particular coalition a interacts with another coalition b. We have an invariant defined in relationship to this decision metric, $Inv^{(0)}(\sigma)$.

For iterated games, we consider chains of games linked together. Consequently our atomic component is not $\Delta^{(0)}$ but $\mathcal{F}^{(0,1)} := \Delta^{(1)} := \{f|f : \Delta^{(0)} \to \Delta^{(0)}\}$, or function spaces over $\Delta^{(0)}$. We have a decision tensor, $\tau : T\mathcal{F}^{(0,1)} \times T\mathcal{F}^{(0,1)} \times T\mathcal{F}^{(0,1)} \to R$, which defines a decision for a coalition a in terms of two different coalitions b and c in adjacent iterations.

Metagames contain games within them. We are thereby interested in the set of games $\mathcal{G}^{(0)} := \{f|f : \Delta^{(0)} \times \Delta^{(0)} \times \Delta^{(0)} \to R\}$. These objects become points in our metagame. In particular, for one-off metagames, we consider a 5-tensor $\kappa : T\mathcal{G}^{(0)5} \to R$. Here the ingredients are: our reference coalition, a, a fixed bridge coalition b (between the game and the metagame), a fixed adjudication coalition h (the house), and a representation of a coalition b in the game, and a representation of a coalition c operating in the metagame. So this makes for five coalitions; consequently, a 5-tensor.

For iterated metagames, we consider chains of metagames linked together. Consequently our atomic component is not $\mathcal{G}^{(0)}$ but $\mathcal{F}^{(1,1)} := \mathcal{G}^{(1)} := \{f|f : \mathcal{G}^{(0)} \to \mathcal{G}^{(0)}\}$, or function space over $\mathcal{G}^{(0)}$. We consider a 7-tensor $\lambda : T\mathcal{F}^{(1,1)7} \to R$. We have here as before our reference a, bridge b, and house h. We also have our coalition representation in the nth iterate of the game b_0, our coalition representation in the $n+1$st iterate b_1, our coalition representation in the mth iterate of the metagame c_0, and our coalition representation in the $m+1$st iterate of the metagame c_1. This makes for 7 coalitions in total, justifying the dimensionality of said tensor.

Then we have a natural way to compute structural coefficients Γ for said 7-tensor, and can compute an information $Inv^{(1)}$. (See [7] and [8] for details as to how this can be done).

If we then realise that the points of a metagame correspond to the information for particular choices of structure for a game, we intuit that our overall information is $Inv^{(1)} \circ Inv^{(0)}$, for $Inv^{(0)}$ the natural information for a one-off game with metric σ; there are four ways to embed this in a metagame, making for four coordinates realised as CW complexes.

Then optimising this information should provide strategies for a coalition of players to execute on said metagame in the best possible way.

Remark. (The 11 tenets). It is probably no coincidence that the dimensionality of an iterated meta-game (11) should be in accordance with the dimensionality of various moral or ethical systems (epitomised say by the ten commandments, together with "the great commandment"). This has potential implications for AI safety as well as the mathematical / rigorous study of ethics.

3.3 Abstraction

Note that we need not stop here; indeed, we can ask the question as to how one should operate in a meta-meta-game.

For metametagames, we are interested in metagames realised as points within them. We are thereby interested in the set of metagames $\mathcal{H}^{(0)} := \{f | f : \mathcal{G}^{(0)} \times \cdots \times \mathcal{G}^{(0)} \to R\}$. However, things are slightly more complicated; note that the first natural metametagame sits on top of a doubly iterated metagame, which sits on top of a doubly iterated game.

Consider $\Delta^{(2,1)} := \{f | f : \Delta^{(1)} \times \Delta^{(1)} \to \Delta^{(1)}\}$ (slightly more elaborate than function-function space over $\Delta^{(0)}$).

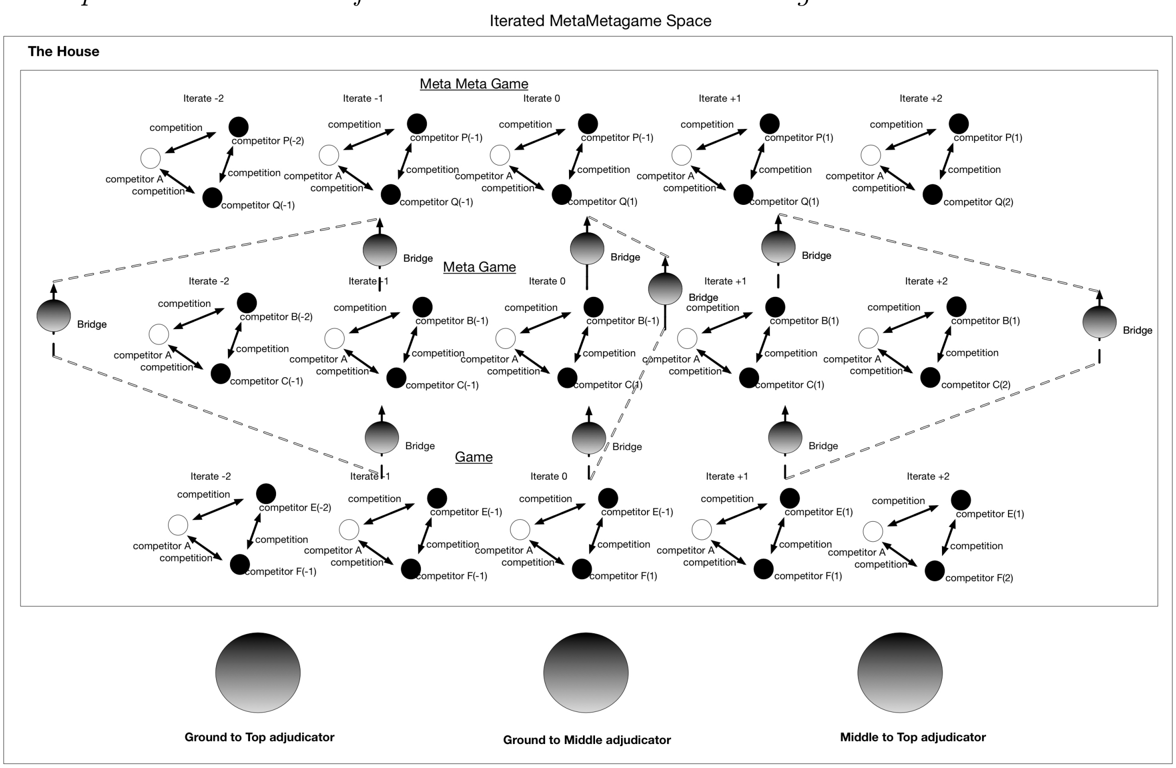
Then with this choice of structure, we are first interested in doubly iterated games $f : \Delta^{(2,1)} \times \cdots \times \Delta^{(2,1)} \to R$ with 11 copies of $\Delta^{(2,1)}$.

On top of our game, we are interested in metagames. We have that:

- One-off metagames are of the form $f : \mathcal{G}^{(0)} \times \cdots \times \mathcal{G}^{(0)} \to R$ with 13 copies, where $\mathcal{G}^{(0)} := \{\phi | \phi : \Delta^{(2,1)} \times \cdots \times \Delta^{(2,1)} \to R\}$ (11 copies).

- We then have one-iterated metagames of the form $f : \mathcal{G}^{(1)17} \to R$, $\mathcal{G}^{(1)} := \{\phi | \phi : \mathcal{G}^{(0)} \to \mathcal{G}^{(0)}\}$.

Example: The structure of a standard iterated meta meta game



- And two-iterated metagames of the form $f : \mathcal{G}^{(2,1)23} \to R$, $\mathcal{G}^{(2,1)} := \{\phi | \phi : \mathcal{G}^{(1)} \times \mathcal{G}^{(1)} \to \mathcal{G}^{(1)}\}$.

Moving finally to metametagames, we have that:

- one-off metametagames are of the form $f : \mathcal{H}^{(0)} \times \cdots \times \mathcal{H}^{(0)} \to R$, with 29 copies, where $\mathcal{H}^{(0)} := \{\phi | \phi : \mathcal{G}^{(2,1)} \times \cdots \times \mathcal{G}^{(2,1)} \to R\}$ (23 copies).
- 1-iterated metametagames are of the form $f : \mathcal{H}^{(1)31} \to R$, where $\mathcal{H}^{(1)} := \{\phi | \phi : \mathcal{H}^{(0)} \to \mathcal{H}^{(0)}\}$.

In this way, in terms of nested layers of iteration, we follow the Fibonacci sequence plus one, working backwards from deepest subtlety (i.e., 1+1 layers for the 0th layer, 1+1 layers for the 1st layer, and 0+1 layer for the 2nd layer). This is important due to the connection with the Golden Ratio. As to proving more rigorously why we would select this form of nesting, I will leave that as an exercise to take up at a later date.

Remark. (Prime number sequences).

If $p : N \to N$ is the prime number sequence, for a metametagame, the next prime by index is:

$$\circ(p; i)(1).$$

- $0 \mapsto \circ(p; 0)(1) = Id(1) = 1$, so index 1 after 19, which is 23.
- $1 \mapsto \circ(p; 1)(1) = p(1) = 2$, so index 2 after 19, which is 29.
- $2 \mapsto \circ(p; 2)(1) = p(2) = 3$, so index 3 after 19, which is 31.

Note for metagames, we have the next prime by index merely as:

$$n \mapsto p \circ p(n)$$

and for games, we have that the next prime by index is:

$$n \mapsto p(n)$$

Structures over a metametagame turn out to be 19-tensors, and we find that the natural information invariant for a meta-meta-game is a 19th order invariant $Inv^{(2)}$. As in [7], we find that strategies are defined in terms of:

$$Inv^{(2)} \circ Inv^{(1)} \circ Inv^{(0)} = 0$$

Due to the nature of the system, we find that $Inv^{(1)}$ is most naturally thought of as a 7×3 or 21st order invariant (multiplicity of the gluing in terms of embedding it in a more complex structure). Similarly, $Inv^{(0)}$ is most naturally thought of as a 3×5 or 15th order invariant, due to the increase in complexity in terms of embedding it in more complex games (doubled multiplicity of the set of possible gluings) (following [7]).

Solving this PDE of order 53 (17th order + 21st order + 15th order) allows us to determine optimal strategies for our system.

Remark. (A puzzle). Should this be an invariant of order 54? That might be more natural. But where would the missing additional derivative come from? Does this come from within a meta-meta-game as an acknowledgement of potential additional abstraction?

In fact, this is a valid question - and leads us to reformulate very slightly what our invariant for a meta-meta-game should be: it should be defined in terms of a tuple consisting of a 31-tensor Ω (the *tail*) with a 1-tensor / 1-form α (the *source*). The role of $\alpha : T\nabla \to R$ is to describe the interaction of the meta-meta-game with the sum total of further levels of abstraction ∇ (i.e., meta-meta-meta-games, meta-meta-meta-meta-games, $\cdots$, etc.). Then we have a 53rd order invariant for interaction with *levels of lower complexity*, and a 1st order invariant for interactions with *levels of higher complexity*, and we multiply these together. This makes for our overall PDE / invariant of order 54:

$$\circ(Inv^{(2)}_{\Omega} \circ Inv^{(1)} \circ Inv^{(0)}(\omega); J_{\alpha}(a)) = 0$$

where $a \in \nabla^{(0)}, \omega \in \Delta^{(2)}$, and $\circ(f; g)$ means that g copies of f are to be composed with each other, i.e. $f \circ \cdots \circ f$, where the number of compositions will generically not be a natural number.

$J_{\alpha}(a)$ in itself could be thought of as 'an entropy':

$$J_{\alpha}(a) := \int_{a \in \nabla} \alpha(a) \partial^{(0)} \alpha(a)$$

as $\partial^{(0)}$, or the logarithmic map, is the inverse of the exponential map on a Riemannian manifold.

Note also that if we take the derivative inside the 1-form α that $S(a) := a\partial a \sim \partial a$, where $\partial := \partial_\nabla$ is the boundary operator on ∇ viewed as a CW complex (or some form of equivalent vector or infinite tensor product of same), and clearly if we define $a *_\nabla b := a \cap b$, then $a \cap \partial a$ is clearly just ∂a, its boundary.

As a final remark before I conclude this component of the discussion, I suggest the following chain of definitions, hearkening back to earlier in the paper:

Definition 1. (First order super-intelligence) (1-Intelligence). *First order super-intelligence*, or *1-Intelligence*, is the capacity to make decisions under uncertainty within a doubly iterated meta-meta-game (hitherto referred to simply as a "game").

In particular, a system displays *(first order) super intelligence* if it can formulate an approach to making decisions in said game under uncertainty, i.e. incomplete information as to the strategies of other players.

Suppose systems A and B both display 1-intelligence (so that benchmarking between the two is meaningful). Then we say that system A is *more intelligent* than system B within a particular game $\mathcal{G}$, if it can formulate more effective strategies to achieve a better payoff within $\mathcal{G}$ when competing against that other system, amongst potentially additional observants.

If this holds for the majority of games that these systems can play with respect to some suitable measure μ on the space of games, then we say that system A is *smarter* or *more capable* than system B with respect to measure μ.

4 Conclusions and a programme for further work

4.1 Summary

Note that this paper was inspired by the observation within [7] that various topological structures seemed to naturally arise in the study of geometry, even if one wished to minimise the introduction of complexity caused by such objects. This led me to ask the question as to what sort of structures one might be able to study if one threw the geometry away, and focused on a systematic study of emergent complexity in topological structures. This paper is an outworking of that.

Indeed, in this paper, I have demonstrated that, rather than having no applications, there are many ways that these lines of thinking can be applied. Indeed, in the as yet nascent field of interpretable machine learning, it is becoming clear that there are deep connections between explanations and equilibria in game theory [17], [23]. Game theory, indeed, is a key application of the theory in this paper, not only in terms of describing the nature of certain games, but also in formulating approaches to compute optimal strategies. Reinforcement learning is another area that benefits from this sort of approach.

We have shown moreover that, in combining ideas from information theory, topology, game theory and number theory, that we can start to intuit relationships between these concepts and various ethical principles and moral frameworks. This in itself is very interesting, and could form the foundation of a more systematic study of quantitative ethics in the context of AI safety, as well as law and governance: namely, how to architect policies and laws to better serve societies in terms of improving stability and success of same, in terms of reactivity to a changing landscape of situations, of risks, and of opportunities.

For instance, consider once again the situation that tends to govern this paper: we have a game, and we have agents that might want to determine an optimal strategy for said game and execute on it. However, in the real world, agents will not have total information, and may not be rational. Therefore, it makes sense to find a way to encourage agents in the aggregate to follow a particular strategy. Call these actions *laws* or *policies.* i.e.

Definition 2. (Policy) A *policy* is a construct out of data relating to some form of summary statistics for a game over a population of agents. The goal of a policy is to encourage agents to behave in ways which are closer to an optimal strategy on the level of a particular set of coalitions of agents.

In particular, one might be interested in looking at expressions such as $g_1 \circ g_0$, or $\circ(g_2 \circ g_1 \circ g_0; h)$, and making these statements about groups or algebras in some logical sort of way, possibly as invariants of same. This I claim would be the logical framework to angle towards in terms of figuring out how to construct policies.

This is the approach I will take in the next paper. However, for now, let us look at something a little bit more whimsical.

4.2 A whimsical look at competition and cooperation between 1-agents

The notions of the preceding analysis have been subject to the assumption that players competing are bayesian, i.e., dominated in decision making capability by a prior and an 'update' of observation. Call these 0-agents. It is useful to ask, though, if we could start to construct 'a game theory for 1-agents', or agents capable of 1-intelligence.

To outline an approach to this, start with the idea of an iterated game, $h : \Delta^{(1)5} \to R$, compared to a one-off game, $f : \Delta^{(0)3} \to R$. Then a logical starting point / springboard to commence the next step in simplification of structure, and exploration of more nebulous concepts, would be to consider the following definition:

Definition 3. ((1,0)- and (1-1)-games) A (one-off) *(1,0)-game* is characterised by a payoff function $f : \Delta^{(1)3} \to R$, i.e. a payoff function $f : \Delta^{(1)} \times \Delta^{(1)} \times \Delta^{(1)} \to R$. An *iterated (1,0)-game* or a *(1,1)-game* is characterised by a payoff function $h : \Delta^{(2)} \times \cdots \times \Delta^{(2)} \to R$, where seven copies of $\Delta^{(2)}$ are taken.

Definition 4. A *(1,n)-game* is a (0,n)-game between *1-agents* (agents capable of 1-intelligence).

Here is one preliminary exploration:

Lemma 4.1. *(Optimal strategies for 1-agents in a (1, 2)-game). The PDE that must needs be solved to determine an optimal strategy for a 1-agent in a (1,2)-game (without gluing), is of order 42.*

Proof. (Sketch).

Consider a metametagame between 1-agents. Then this is characterised by a function

$$p^{(1,0)} : \mathcal{J}^{(1,1)31} \to R$$

where

$$\mathcal{J}^{(1,1)} = \{\phi | \phi : \mathcal{J}^{(1,0)} \to \mathcal{J}^{(1,0)}\}$$

and

$$\mathcal{J}^{(1,0)} := \{\lambda | \lambda : \mathcal{F}^{(1,2)} \to R\}$$

where

$$\mathcal{F}^{(1,n+1)} := \{\phi | \phi : \mathcal{F}^{(1,n)} \to \mathcal{F}^{(1,n)}\}$$

and

$$\mathcal{F}^{(1,0)} := \{h | h : \Delta^{(2)11} \to R\}$$

Then we have a 31st order invariant $Inv^{(1,2)}$ for the metametagame. For the metagame, we have a 7th order invariant, and for the game, we have a 3rd order invariant. We also have a 1-form α with an associated entropy.

Putting these together, we determine that the PDE which needs must be solved to determine an appropriate strategy *without gluing* in a (1,2)-game between 1-agents is of order 42.

Remark. Now recall that coalitions of 0-agents in a metametagame have a strategy governed by a PDE of order 54. So why consider 1-agents playing an analogous metametagame (albeit on a higher level) without gluing? The trick is that, for 0-agents, individual 1-agents, or coalitions of same, are indistinguishable. Gluing only occurs / makes sense in cooperative games, where "CW complexes are non-trivial". In particular, a 0-agent cannot distinguish between multiple 1-agents, or multiple 0-avatars of a single 1-agent. Hence the strategy for a 1-agent playing a similar metametagame is, from the point of view of the 0-agents, effectively of order 42.

□

References

[1] A Adem, D Cohen, and FR Cohen. On representations and k-theory of the braid groups. *arXiv preprint math/0110138*, 2001.

[2] Minhyong Kim et al. Unramified arithmetic chern-simons invariants. *https://arxiv.org/abs/1705.07110*, 2017.

[3] Chris Goddard. Advanced topics in information dynamics. *viXra preprint viXra:1002.0033*, 2010.

[4] Chris Goddard. Towards a general theory of statistical inference. *viXra preprint viXra:1506.0067*, 2015.

[5] Chris Goddard. Optimisation of dynamical systems subject to meta-rules. *viXra preprint viXra:1604.0148*, 2016.

[6] Chris Goddard. A criterion arising from explorations pertaining to the oesterle-masser conjecture. *viXra preprint viXra:1705.0115*, 2017.

[7] Chris Goddard. Soliton solutions to the dynamics of space filling curves. *viXra preprint viXra:1711.0459*, 2017.

[8] Christopher Goddard. An information theoretic approach to game theory, ı. *viXra preprint viXra:1810.0289*, 2018.

[9] Christopher John Goddard. A treatise on information geometry. *arXiv preprint arXiv:1802.06178*, 2018.

[10] Mikhail Kapranov and Vladimir Voevodsky. Braided monoidal 2-categories and manin-schechtman higher braid groups. *Journal of Pure and Applied Algebra*, 92(3):241–267, 1994.

[11] Minhyong Kim. Arithmetic chern-simons theory i. *arXiv preprint arXiv:1510.05818*, 2015.

[12] Minhyong Kim, Hee-Joong Chung, Dohyeong Kim, Georgios Pappas, Jeehoon Park, and Hwajong Yoo. Abelian arithmetic chern–simons theory and arithmetic linking numbers. *International Mathematics Research Notices*, 2017.

[13] Minhyong Kim, Hee-Joong Chung, Dohyeong Kim, Jeehoon Park, and Hwajong Yoo. Arithmetic chern-simons theory ii. *arXiv preprint arXiv:1609.03012*, 2016.

[14] Ki Hyoung Ko and Hyo Won Park. Characteristics of graph braid groups. *Discrete & Computational Geometry*, 48(4):915–963, 2012.

[15] Hisatoshi Kodani, Masanori Morishita, and Yuji Terashima. Arithmetic topology in ihara theory. *arXiv preprint arXiv:1608.07926*, 2016.

[16] Tom Leinster. Higher operads, higher categories. *arXiv preprint math/0305049*, 2003.

[17] Scott Lundberg and Su-In Lee. A unified approach to interpreting model predictions. *arXiv preprint arXiv:1705.07874*, 2017.

[18] Barry Mazur. Remarks on the alexander polynomial. *Unpublished*, 1963.

[19] Barry Mazur. Notes on étale cohomology of number fields. In *Annales scientifiques de l'École Normale Supérieure*, number 4 in 6, pages 521–552. Elsevier, 1973.

[20] John Milnor and James D Stasheff. *Characteristic Classes.(AM-76)*, volume 76. Princeton university press, 2016.

[21] Masanori Morishita. *Knots and primes: an introduction to arithmetic topology*. Springer Science & Business Media, 2011.

[22] Igor Nikolaev. Remark on arithmetic topology. *arXiv preprint arXiv:1706.06398*, 2017.

[23] Roberto Serrano. Cooperative games: Core and shapley value. Technical report, Working Paper, Brown University, Department of Economics, 2007.

[24] Ross Street. Higher categories, strings, cubes and simplex equations. *Applied Categorical Structures*, 3(1):29–77, 1995.

[25] Jean-Luc Thiffeault and Matthew D Finn. Topology, braids, and mixing in fluids. *arXiv preprint nlin/0603003*, 2006.

Links

[L1] Chao and Charmaine. Knots and primes (lecture notes). http://www.math.columbia.edu/ chaoli/tutorial2012/.

[L2] Jordan S. Ellenberg and Daniel Rayor Hast. Rational points on solvable curves over q via non-abelian chabauty. https://quomodocumque.wordpress.com/2017/06/05/rational-points-on-solvable-curves-over-q-via-non-abelian-chabauty-with-daniel-hast/.

[L3] Minhyong Kim. Arithmetic topological quantum field theory? (talk). https://simonsfoundation.s3.amazonaws.com/share/mps/conferences/2017_Conference_on_Number_Theory_Geometry_Moonshine_and_Strings/Kim.pdf.

[L4] Minhyong Kim. The unipotent albanese map and selmer varieties for curves. http://www.kurims.kyoto-u.ac.jp/ prims/pdf/45-1/45-1-4.pdf.

[L5] Nilay Kumar. Characteristic classes ii.
http://www.math.northwestern.edu/ nilay/pdf/characteristicclasses.pdf.

[L6] Wikipedia. Arithmetic topology.
https://en.wikipedia.org/wiki/Arithmetic_topology.

[L7] Wikipedia. Braid theory. https://en.m.wikipedia.org/wiki/Braid_theory.

[L8] Wikipedia. Knot theory. https://en.m.wikipedia.org/wiki/Knot_theory.