

# Model for multi-agent Optimisation

Ambrose Yim (University of Oxford)

2 May 2018

## Introduction

Let there be  $N$  agents indexed by  $i \in \mathcal{I}$  working together on a project proposal. Each agent has a (finite) set of actions  $A_i$  which they could employ to modify particular contents of the project proposal. The state of the entire project proposal sits in the product space of all actions  $A = A_1 \times \dots \times A_N$ . We associate to each agent a utility function,  $u : \mathcal{I} \times A \rightarrow \mathbb{R}$ , which we can interpret as the agent's happiness at a given state.

While the objective of an individual agent is to selfishly maximise their own utility, the project only goes forward when all of the agents are 'sufficiently happy': If any one agent is too unhappy, the project cannot proceed.

Firstly, we define the **collective optimisation problem** (the Co-Op) where where all agents sit together, *pool* their information - i.e. their utility functions - and *altruistically* maximise their **collective benefit function**  $U(a) = \min_{i \in \mathcal{I}} u_i(a)$ , so that no agent is 'left behind' in the name of 'progress' <sup>1</sup>:

$$\max_{a \in A} \min_{i \in \mathcal{I}} u_i(a) = U^*$$

The optimum (optima) of this problem is  $U^* = u_{i^*}(a^*)$ . There are good reasons to choose this philosophy: quite often there is no actual global utility function to optimise over agents. Multi-agent algorithms often optimise the sum of local utility functions which is an arbitrary choice. A bad reason to adopt this philosophy is when there is a sensible definition of global utility function.

Suppose each agent acts selfishly and wants to push the project towards a state that maximises their own utility  $\max_{a \in A} u_i(a)$ . Then if all agents optimising selfishly can all do better than  $U^*$ , then they would not end up in the state  $a^*$  that optimises Co-Op. In other words, the condition that selfish agents fail to 'Co-Op' is

$$\max_{a \in A} \min_{i \in \mathcal{I}} u_i(a) < \min_{i \in \mathcal{I}} \max_{a \in A} u_i(a)$$

However there is this famous inequality, well known in game theory and optimisation circles:

$$\max_{a \in A} \min_{i \in \mathcal{I}} u_i(a) \leq \min_{i \in \mathcal{I}} \max_{a \in A} u_i(a)$$

with stringent conditions on  $\mathcal{I}$ ,  $A$  and  $u$  for equality to hold <sup>2</sup>. So generically, selfish actions cannot lead to the 'Co-op' communist utopia.

**Question:** Can agents strike a balance between altruistic and selfish intentions?

---

<sup>1</sup>Aside: if  $u$  is a binary function, where an agent gives assent to a project proposal if  $u = 1$  and vetoes the proposal if  $u = 0$ , the Co-Op optimum is simply the condition that  $u_i(a^*) = 1 \quad \forall i \in \mathcal{I}$ .

<sup>2</sup>There is a theorem in convex optimisation - if  $\mathcal{I}$ ,  $A$  are convex subsets of some field (e.g. rationals, reals) and  $u$  is a convex function then equality is ensured.

## Model for Collaborative Process

However there is more finesse to this argument. The assumption that a selfish agent can maximise over all states  $A$  is unrealistic - that is what each agent wants; whether it can actually do so is another matter entirely, restricted by its restricted influence on the project proposal. In complex problems (like chess) it is also unlikely that it would be computationally efficient to pre-compute  $u$  for all states of the system - usually we can only evaluate  $u$  *during* the game.

This points towards the temporal nature of real life organisational optimisation problems.

### Coalitions and Coalition Operators

Suppose in an optimisation problem we start in some state  $a_0 \in A$  and evolves in discrete time steps  $a_0 \rightarrow a_1 \rightarrow \dots \rightarrow a_T$ . Let  $K = \{C \subseteq \mathcal{I}\}$  be the collection of all subsets of agents and denote  $A(C)$  to be the actions available to agents in  $C$ . At any moment  $t$  a subset of agents  $C_t$  come together as a **coalition** to solve the the problem

$$a_t = \arg \max_{a \in A(C_t)} \min_{i \in C_t} u_i(a|a_{t-1}) = \mathcal{C}_t a_{t-1}.$$

We call  $\mathcal{C} : A \rightarrow A$  a **coalition operator**.

N.B. in a coalition, members

- Coordinate their actions (in max operation)
- Share information about their utility function (in min operation)
- The state space they could explore is restricted by its immediate predecessor  $a_{t-1}$
- The coalition updates the system to a new state where the worse off member of the coalition has a greater utility

*Points to Explore:*

- What are the constraints on the coalitions that we are allowed to form?
- There is scope for the coalition to be more altruistic/more knowledgeable, by minimising  $u_i$  over a superset of  $C_t$  (i.e. agents have limited actions but better information and less selfish)
- How does the problem restrict the image of coalition operators?

### Collaborative Process

Thus in our framework we can define a **collaborative process**, which is prescribed by a tuple of a set of agents, actions associated with each agent, a utility function, an initial state and a *time-ordered* set of coalitions  $\langle \mathcal{I}, A, u, a_0, K \rangle$  from which a chain sequence of states is evolved

$$a_0 \xrightarrow{\mathcal{C}_1} a_1 \xrightarrow{\mathcal{C}_2} \dots \xrightarrow{\mathcal{C}_T} a_T,$$

where

$$a_t = \mathcal{C}_t \mathcal{C}_{t-1} \dots \mathcal{C}_1 a_0.$$

To abstract the problem in a more succinct manner, we associate to every ordered set of coalitions a **sequence operator**  $\mathcal{K} : A \rightarrow A$ , which is a map from an initial state to a final state of the sequence  $\mathcal{K} = \mathcal{C}_T \mathcal{C}_{T-1} \dots \mathcal{C}_1$ . We observe that the set of sequence operators  $\Gamma$  is a *semigroup* finitely generated by coalition operators  $\mathcal{C}^3$ .

*Points to Explore:*

---

<sup>3</sup>that is,  $\Gamma$  is a ‘group’ without identity and inverses, and every member can be obtained by multiplying together some  $\mathcal{C}$ ’s.

- Is there something special about the algebra of  $\Gamma$ ?
- How could we exploit the structure of the problem (namely, the utility matrix  $u$ ) to design a sequence  $K$  that maximises  $U(a_T)$  with as few time steps  $T$  as possible, or make sure  $U(a_t)$  always decreases? e.g. select certain members of  $C_t$  to carry forward to  $C_{t+1}$  under some criteria?
- Is it possible to actually design  $K$  that decreases  $U$  for a general problem?

## Fixed Points

A **fixed point** of an operator  $\mathcal{K}$  is an element  $\alpha \in A$  if  $\mathcal{K}\alpha = \alpha$ . Moreover it is an **attractive fixed point** if starting from an arbitrary element  $a_0$  in some neighbourhood  $B \subseteq A$  of  $\alpha$  the sequence  $\{a_t \mid a_{t+1} = \mathcal{K}a_t\}$  converges to  $\alpha$ .  $N$  is called the attracting set of  $\alpha$ .

In real life terms, if we can find  $\mathcal{K}$  which has an attractive fixed point, we just need to loop *any* project proposal  $a_0 \in B \subseteq A$  round the sequence over and over again and we are guaranteed to reach a stable agreement.

We could also find, for each  $a$ , the set of sequence operators that maps  $a \mapsto a$ :  $\mathcal{O}_a = \{\mathcal{K} \in \mathcal{O} \mid \mathcal{K}a = a\}$ . This quantifies the stability of each state against a coalition. It is worth remarking that it is quite possible that  $\mathcal{O}_{a^*} \neq \Gamma$  i.e. for the optimum of Co-Op is *not* stable against coalition decision making; in the worst case,  $\mathcal{O}_{a^*} = \{\mathcal{I}\}$ .

We can flesh out this landscape of fixed points and attracting sets with two approaches. One is to construct a metric space  $(A, d)$ , i.e. impose a measure of distance between states; the other is to represent our system of states and operators in terms of linear algebra.

## Metric space approach

If we have a (non-empty, complete)<sup>4</sup> metric space  $(A, d)$ , we can make use of the Banach fixed point theorem. If  $\mathcal{K}$  is a contraction, i.e.  $\exists q \in [0, 1)$  s.t.  $\forall \alpha, \beta \in A, d(\mathcal{K}\alpha, \mathcal{K}\beta) \leq qd(\alpha, \beta)$ , there exists a *unique* attractive fixed point  $\mathcal{K}\alpha^* = \alpha^*$ .

What remains is to choose a metric that sensibly models how. There are two obvious choices for the metric  $d : A \times A \rightarrow \mathbb{R}_+$ . The first one is to embed  $A \hookrightarrow \mathbb{R}^N$  ( $N$  being the number of agents) using the utility function as a coordinate map:  $a_i = u_i(a)$ . This induces a Euclidean metric on  $A$ . However, since the ideal situation is to solve the ‘Co-Op’ problem, we can make a legitimate simplification and choose  $(a, a') \xrightarrow{d} |U(a) - U(a')|$ , i.e. the difference between two states is their relative performance in optimising ‘Co-Op’.

## Linear Algebra Approach

We define  $V(A)$ , a real vector space taking elements of  $A$  as its basis  $\{\varphi_a\}_{a \in A}$ . Since  $A = A_1 \times \dots \times A_N$ , our vector space is a tensor product space, allowing for a compact representation:  $V(A_1 \times \dots \times A_N) = V(A_1) \otimes \dots \otimes V(A_N)$  where  $\otimes$  is the *kronecker product*.  $\Gamma$  induces a set of *linear* operators  $L$ ; if  $a \xrightarrow{\mathcal{K}} a'$ ,

---

<sup>4</sup>Since  $A$  is finite we are fine.

there is a corresponding linear operator  $l(\mathcal{K}) \in L$  such that  $\varphi_a \xrightarrow{l(\mathcal{K})} \varphi_{a'}$ . We have the commutative diagram:

$$\begin{array}{ccc} A & \xrightarrow{\mathcal{K}} & A \\ \downarrow & & \downarrow \\ V & \xrightarrow{l(\mathcal{K})} & V \end{array}$$

The action of the linear operator  $l(\mathcal{K})$  can be represented by a matrix  $\mathcal{K}$  (overloading notation since it is becoming ridiculously cumbersome?): entries  $\mathcal{K}_{a'a} = 1$  if  $\varphi_a \xrightarrow{l(\mathcal{K})} \varphi_{a'}$ , and  $\mathcal{K}_{a'a} = 0$  if otherwise. In other words, the  $a'^{\text{th}}$  row in the  $a^{\text{th}}$  column is one if and only if the corresponding sequence operator sends state  $a$  to  $a'$ , otherwise it is zero.

We can still decompose  $\mathcal{K}$  into products of generators  $\mathcal{C}$ . Needless to say  $L$  is also a finitely represented semigroup closed under matrix multiplication.

A fixed point of operator  $\mathcal{K}$  is then simply the eigenvector of the matrix  $\mathcal{K}$  with eigenvalue one. We can take a matrix and compute its singular/eigenvalues for a stability analysis of that operator.

*Comment:* Given the particular structure of the matrices, analysis should be tractable...

*Points to Explore:*

- Decomposition of operators into tensor products: what is the significance? look at high order svd.

## Summary Remarks

The really hard but interesting question is the relationship between the utilities of the agents  $u$  and the optimisation performance of sequences. Specifically, how *correlations* between decisions made by agents affect the *algebra* between coalition operators that generate the sequences. Some nice results should happen if agent decisions are uncorrelated/positively correlated/ anti-correlated. (We need some measure of correlation)

We expect computation and simulation to be expensive; complexity of computation mentioned so far is exponential in the number of states.

If we have time, We can hop onto the intellectual gravy train of machine learning and pretend some reinforced learning algorithm adaptively learning sequences is a good model of human behaviour. Further work can also be done on generalising the coalition operator e.g. do agents retain *memory* from previous coalitions?