ESSAI-2024
Self-Governing Multi-Agent Systems
L8/10: Consensus

Jeremy Pitt and **Asimina Mertzani**

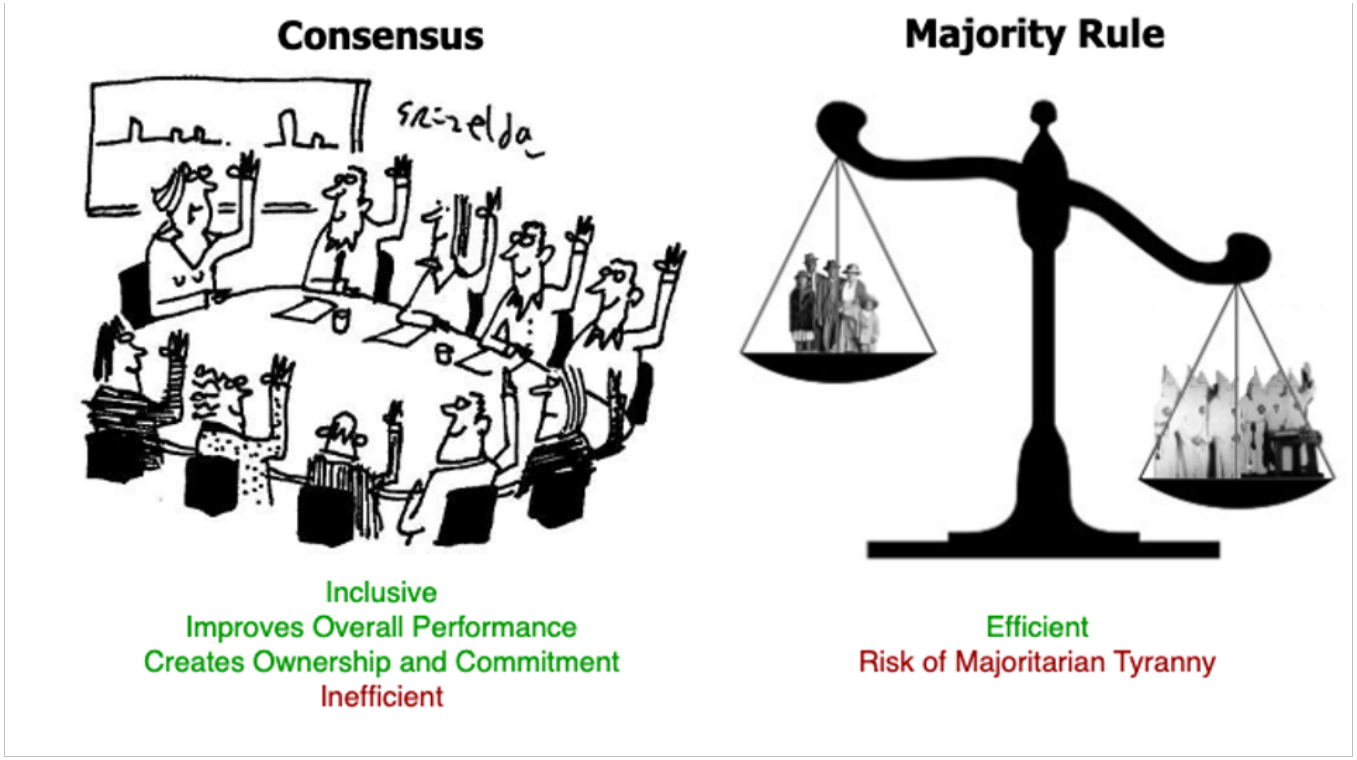
Department of Electrical and Electronic Engineering

Imperial College London

Aims and Objectives

- Aims
  - learn a mechanism for consensus formation used by classical Athenian assemblies, called thoryvos
  - learn how to extend Q-learning algorithm to develop a formal model of thoryvos, called Θ-learning
- Objectives
  - bring together political sciences with machine learning to facilitate consensus formation in SGMAS

Consensus vs Majority Rule



- Dissent
  - **Obstructive dissent:** when groups block proposals for the common good in order to protect their own self-interest
  - **Productive dissent:** for reformation when current practices diverge from core values

Majoritarian Voting – Plurality

- Suppose: two alternative candidates
- Suppose: set of voters create *profile* $\langle 0, 1, -1, \ldots, 1 \rangle$
- Some procedural requirements ('nice' requirements)
  - Universality: all possible profiles are admissible inputs
  - Anonymity: permutations produces the same result
  - Neutrality: invert profile, invert result
  - Positive responsiveness: if some voters change their vote in favour of one alternative, the result does not change in favour of the other
- Theorem (May, 1952)
  - An aggregation rule satisfies universal domain, anonymity, neutrality and positive responsiveness if and only if it is a majority rule
- What could possibly go wrong?
- Just add a third alternative...

Condorcet Winner and Condorcet Loser

- Assume there are *three* candidates, $C = \{a, b, c\}$
  - Then a voter's preference (rank order) can be any one of six possible linear orderings over C

Preference	P_1	P_2	P_3	P_4	P_5	P_6
	a	a	b	b	c	c
	b	c	a	c	a	b
	c	b	c	a	b	a
# voters	n_1	n_2	n_3	n_4	n_5	n_6

- Majority relation $\succ_m$ which ranks the candidates according to how they fare in one-to-one comparisons
- The **Condorcet Winner** is the candidate that is maximal in the majority relation $\succ_m$, i.e. it wins more one-to-one comparisons that any other candidate
- Condorcet's Paradox: even if each voter's preference ordering is transitive, the majority ordering might not be

Implications

- Plurality (relative majority) aggregation rule is simple, intuitive, plausible and most likely to produce a winner
- But (when there are more than 2 alternatives)
  - Can elect the Condorcet Loser
  - Loses information
- So use an alternative method

Voting Methods

- Methods
  - Plurality, Runoff, Borda Count, Instant Runoff, Approval
  - Copeland Scoring
  - D'Hondt System
- But
  - Different voting methods can produce different results *from the same votes*
  - More complex (to understand and to compute), can give unintended consequences
  - Voting methods are susceptible to strategic manipulation
- From Voting to judgements

D'Hondt System

- Used in European Parliament Elections in UK
  - Multiple winner election in a constituency
  - Each party submits a ranked list of candidates for n winners
  - Each voter votes for a *party* (not a particular candidate)
  - Method
    - Divide votes-for-party by (number-of-winners-for-party + 1)
    - Party with most votes gets 1 winner
    - Repeat until n winners
  - The pursuit of fairness may have unintended consequences...

Manipulation: Example from Pliny (more or less)

- Death of a Roman Consul: the slaves stand accused of his killing. But:
  - Consul committed suicide $\rightsquigarrow$ acquittal
  - Slaves assisted suicide $\rightsquigarrow$ banishment
  - Slaves murdered $\rightsquigarrow$ death
- The Senate has to decide
  - Acquittal and death are opposites
    - Think guilty (death): banishment is 'preferable' to acquittal
    - Think innocent (acquittal): banishment is 'preferable' to death
  - Doubt (banishment): acquittal is 'preferable' to death
- Three factions in the Senate
  - Faction D: 37%: Death $\succ$ Banishment $\succ$ Acquittal
  - Faction B: 35%: Banishment $\succ$ Acquittal $\succ$ Death
  - Faction A: 28%: Acquittal $\succ$ Banishment $\succ$ Death

So What Can Happen?

- If you are in Faction D: insist on a ternary vote, plurality wins
- If you are in Faction B: insist on pairwise comparison votes, most winning comparisons wins
  - A vs. B: B wins 72 to 28
  - A vs. D: A wins 63 to 37
  - B vs. D: B wins 63 to 37
- If you are in Faction A
  - Realise there are two votes: one for innocence or guilt, and if guilty, another for punishment (death or banishment)
  - So arrange for two votes, but:
    - Have the punishment vote *first*, and
    - Vote *against* your own preference
  - Then: Death beats Banishment in the first vote (65 to 35)
  - Acquittal beats Death in the second vote (63 to 37)

Judgement Aggregation

- Three people share a flat
- Mutually agreed a set of conventional rules, Including
  - If the fridge is dirty, and the cooker is dirty, then the kitchen should be cleaned
- In their judgements:
  - One person thinks the fridge is dirty, but the cooker is clean
  - Another person: the cooker is dirty, but the fridge is clean
  - The third person: both the fridge and the cooker are dirty
- Should the kitchen be cleaned?

	p (fridge dirty)	q (cooker dirty)	$p \wedge q$
Hejhog1	**true**	**false**	**false**
Hejhog2	**false**	**true**	**false**
Hejhog3	**true**	**true**	**false**
Majority	**true**	**true**	**false**

Social Choice Theory

- No end of profound and interesting results
- No end of paradoxes
  - Arrow's Impossibility Theorem
- What to conclude
  - "Dozens of possible voting methods have been devised, ranging from the imperfect to the abysmal" (Monbiot, 2017)
  - "Most systems are not going to work badly all of the time. All I proved is that all can work badly at times" (Arrow, 2008)
  - Maybe we are looking in the wrong place

Motivation

We want a **sustainable mechanism** for forming collective agreements (i.e. decisions) in socio-technical systems.

What can we do?

Let's draw some (more) inspiration from history...

Classical Athenian deliberative assemblies seemed to be quite effective at that.

So let's look into what were they doing.

Historical Perspective

Deliberative assemblies in classical Athens sought **consensus** but "didn't mind" using majority rule.

What was their edge?

They were using a process called 'Thoryvos'.

Thoryvos

Thoryvos: a process to **detect emergent consensus** in the form of **persistent general agreement** as a prelude to a vote on options (Canevaro 2018).

That general agreement was inferred from the **vocal expressions** of the citizens (i.e. agents).

Thoryvos: The Process

- Speakers were **proposing policy** options to the citizens.
- Citizens were **expressing vocally their opinion** and formed thorybos (cheering, heckling, shouting or muttering).
- Proedroi (i.e. moderators) **selected** policy options to put to a vote **based on the vocal expressions**.

How to Formalise it?

From Consensus Formation in classical Athens to Markov Decision Processes (MDP)

- We can **extract interesting aspects of thorybos** and use them to design (multi-agent) systems for **sustainable self-governance**.
- We can **feed** their **vocal expression** of agreement (or otherwise) on a policy **into a learning process**.
- We propose to model consensus formation in a deliberative assembly using thorybos as a Markov Decision Process (MDP) and use Reinforcement Learning (RL) to **learn how to reach general agreements on diverse preferences with minimal compromises**.

Problem Specification

- Consensus through democratic deliberation.
- Aim: mechanism for learning how to reach agreements on the policies and how to maintain those agreements, regardless of what those policies are.
- Examples of use: startups, cohousing projects, community energy grids

The purpose is **not** *to identify the* **optimal approach**, *but to abstract from it, and propose a* **mechanism** *for reaching consensus and deliberating about the process of reaching consensus.*

Reinforcement Learning Formalisation (Briefly)

- states: collective noise (e.g. thoryvos)
- actions: policies

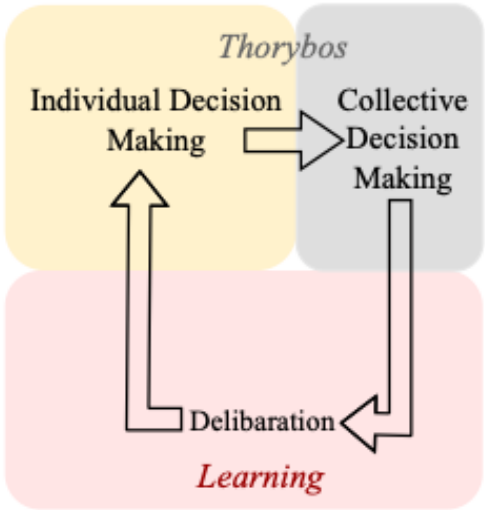
...but agents have preferences (pr) on policies...

- reward: tries to balance out individual with collective good

Θ-Learning: Overview

Two-phased cyclic process comprising by:

- Thorybos: a process of decision making
- Learning: a process of deliberation about the process of decision making

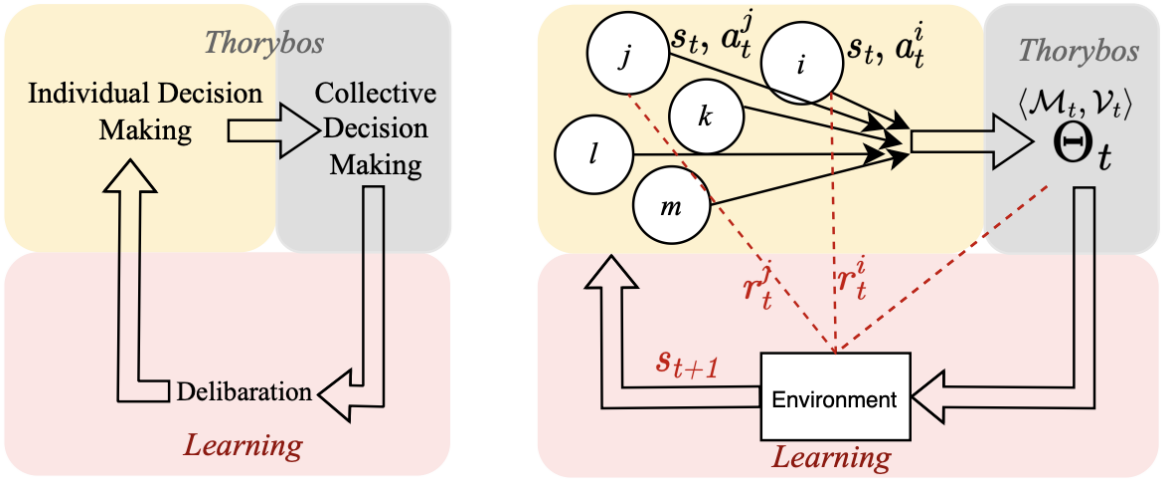


...Yet another hybrid system...

Θ-Learning: More

Two-phased cyclic process comprising:

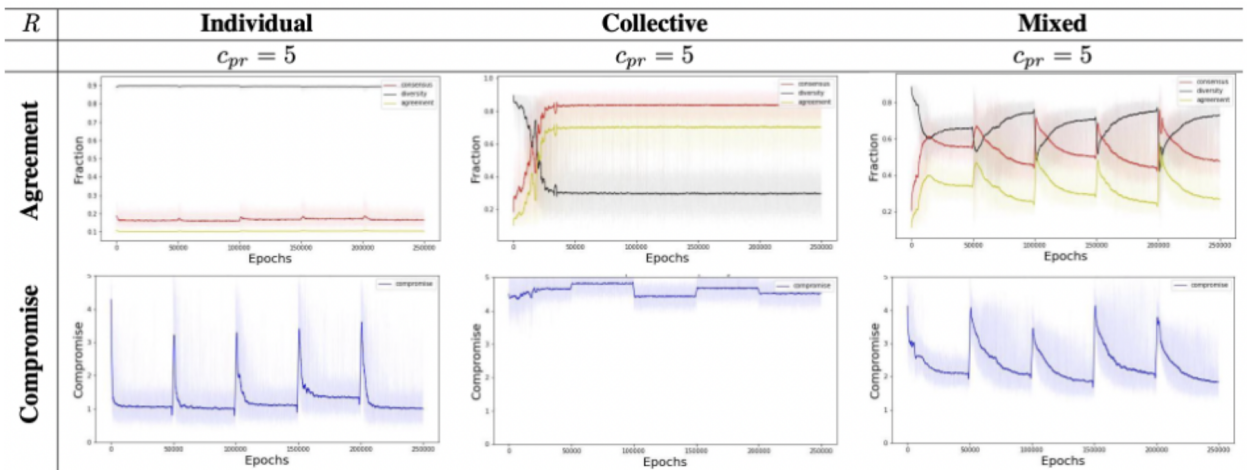
- Thorybos: a process of decision making
  (corresponds to the policy supported by the most, and the number of individuals supporting that policy)
- Learning: a process of deliberation about the process of decision making



Experimental Framework

- Aim: Explore whether Θ-learning can be effective in solving problems of collective decision making.
- Varying Experimental Conditions:
  - Learning Objective
    - Individual: Reinforces individual rationality (aiming for **minimising compromises**)
    - Collective: Reinforces collective contribution (aiming for reaching a **collective agreement)**
    - Mixed: Combines both (aiming for reaching **agreement** while maintaining meaningful dissent)
  - Rate of Change of Preferences
    - static
    - dynamic with different speed of change

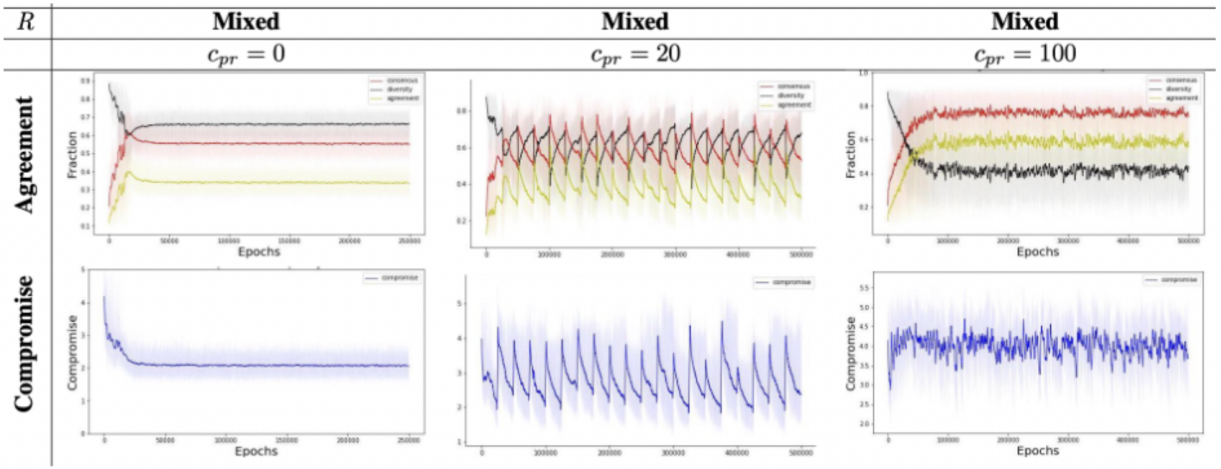
Experimental Results - Effect of Learning Objective



-Red Line: amount of the population proposing the winning policy
-Black Line: dissent expressed
-Yellow Line: agreement reached
-Blue Line: compromises made

- The **objective** of agents, determined by their reward function, **strongly relates** with whether agents manage to reach a **general agreement** and the **compromises** of agents.
- When agents' objective is 'Mixed', Θ-learning provides a mechanism for sustainable self-governance **balancing** out tensions between **consent and compromise**.
- When agents' objective is 'Collective', Θ-learning constitutes a mechanism for **consensus formation**.

Experimental Results - Effect of Rate of Change



-Red Line: amount of the population proposing the winning policy
-Black Line: dissent expressed
-Yellow Line: agreement reached
-Blue Line: compromises made

- The **speed of change** of the population is **correlated** with the form of **collective agreement** and the level of **compromise**.
- The **immediate reaction** to a change is to prioritise processes that reassure **stability** in the short-term, i.e. reaching an **agreement** regardless of the compromises.

Summary of Experiments

- Depending on the learning objective, **different forms of agreement** and **levels of compromise** can emerge, producing different outcomes in terms of **stability**.
- When agents act **individually**, **dissent is expressed** but this leads to **instability**.
- When agents act **collectively**, a **general agreement** is formed but **dissent is suppressed**, which might lead to **stagnation** due to lack of diversity.
- When **individual and collective objectives** are combined, the group forms a **persistent general agreement** which leads to sustainable (quasi-stable) self-governance.

Summary

- The synthesis of thorybos with Q-learning provides a fit-for-purpose algorithm for **sustainable collective self-governance** through meaningful democratic deliberation.
- **Legitimate consent** produces **compromises** which turn out to be useful **conceptual resources** used in future negotiation.
- **Legitimate dissent** extracts a **productive** signal out of what might otherwise be regarded as **distracting 'noise'**.
- **Sustainable** self-governance through democratic participation in deliberation **requires combining both**.
- **Majoritarian** decision-making is **acceptable** if the **underlying principle is consensus** reached through democratic deliberation (Canevaro, 2018).

Takeaway

The analysis of classical documents can lead to the development of technology that enables a group to achieve sustainable self-governance.

- While in theory **no voting procedure can satisfy all the fairness criteria** at the same time, this means the challenge is to **design a voting procedure that minimizes the likelihood of an unfair outcome**
- There are other questions to ask: e.g., it **not just a matter of how someone gets elected** to a position of authority, but **why** they want it, and **what** they do with it if they get it
- **Preferences are not a timeless, infallible and unquestionable product of votes**