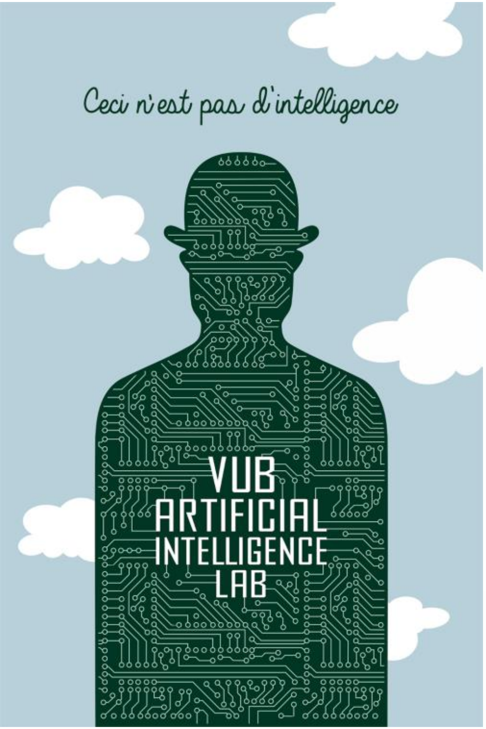


When Multiple Agents Care About More than One Objective

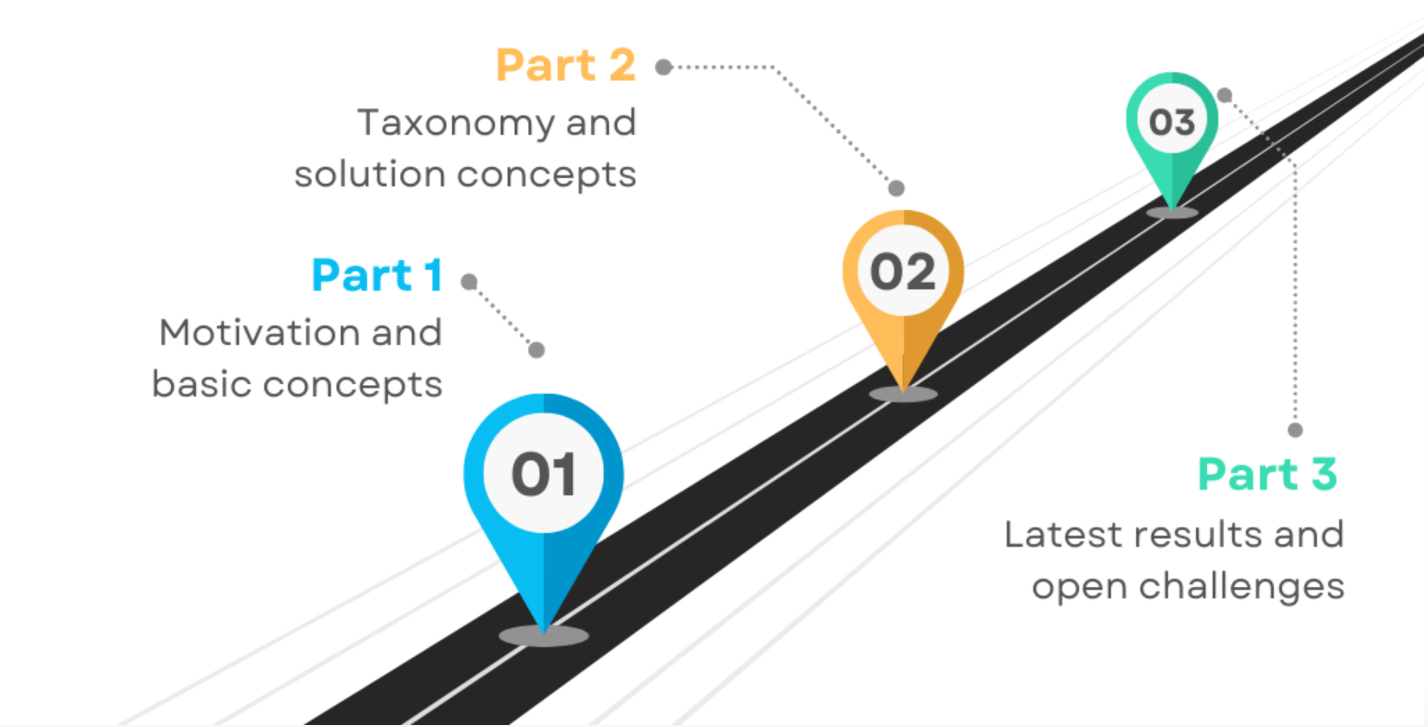
Diederik M. Roijers and Roxana Rădulescu

IJCAI Tutorial, Vienna, 2022

VUB ARTIFICIAL INTELLIGENCE RESEARCH GROUP

@aibrussels

Tutorial Roadmap



Part 1 - Multi-objective decision making in multi-agent systems

Motivation and basic concepts

Going to the conference

Two players

- rewards are public
- utility is private

MONFG

Why hard?

	Taxi	Tram	Walking
Taxi	(10€, 5min); (10€, 5min)	(20€, 5min); (2€, 15min)	(20€, 5min); (0€, 35min)
Tram	(2€, 15min); (20€, 5min)	(2€, 15min); (2€, 15min)	(2€, 15min); (0€, 35min)
Walking	(0€, 35min); (20€, 5min)	(0€, 35min); (2€, 15min)	(0€, 35min); (0€, 35min)

Why?

Multiple objectives

Because life is not simple

- What are your objectives for your current research project?
  - Publishing asap?
  - Quality of conference/journal?
  - Collaboration potential?
  - Flag-posting?
  - Increasing funding potential?
  - Finishing your PhD?

Because life really is not simple

- What are your objectives for your current research project?
  - Publishing asap?
  - Quality of conference/journal?
  - Collaboration potential?
  - Flag-posting?
  - Increasing funding potential?
  - Finishing your PhD?

- How about your co-authors?

Multiple objectives!

- Most decision problems have multiple objectives

- Cannot scalarise a priori
  - Unknown, uncertain, or private utility
  - Non-linear utility
  - Changeable preferences/utility
  - Adjustability
  - Explainability for oversight and review purposes

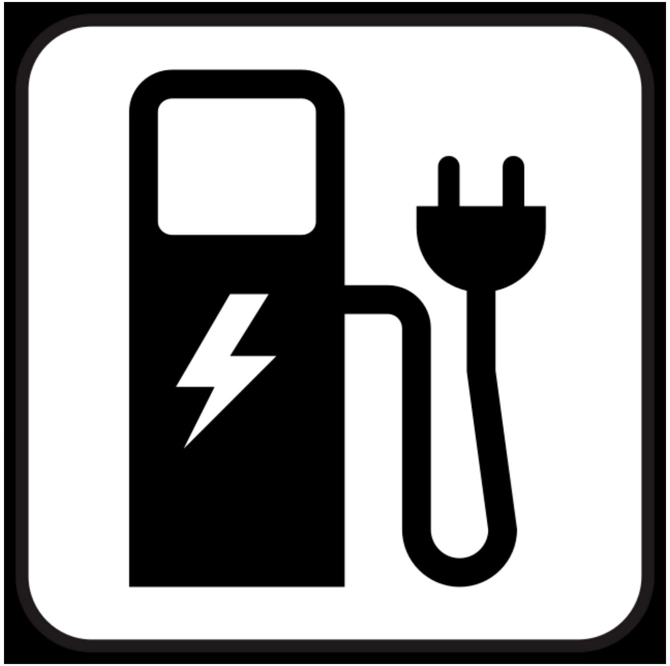
- To scalarise is to throw away information

More and more MO

- AI has ever increasing impact on people's lives
- Ethical aspects more important
  - Human-aligned AI is a multi-objective problem [Vamplew et al., 2018]
- Explainability more important
  - Legal frameworks incoming
- Environmental concerns

Example: electric vehicle charging

- meeting demands
- minimising costs
- preventing grid overloads



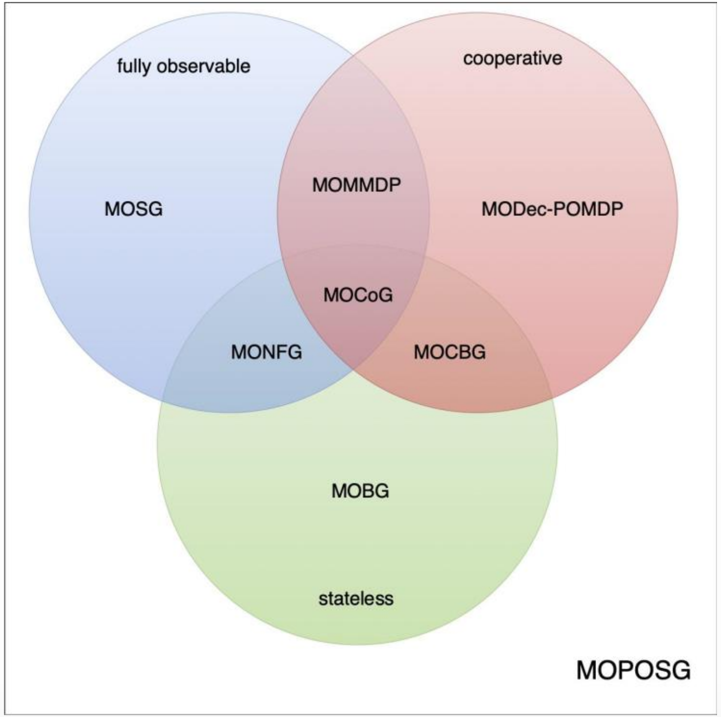
Modelling and dealing w/

Multiple objectives

User utility is central to modelling

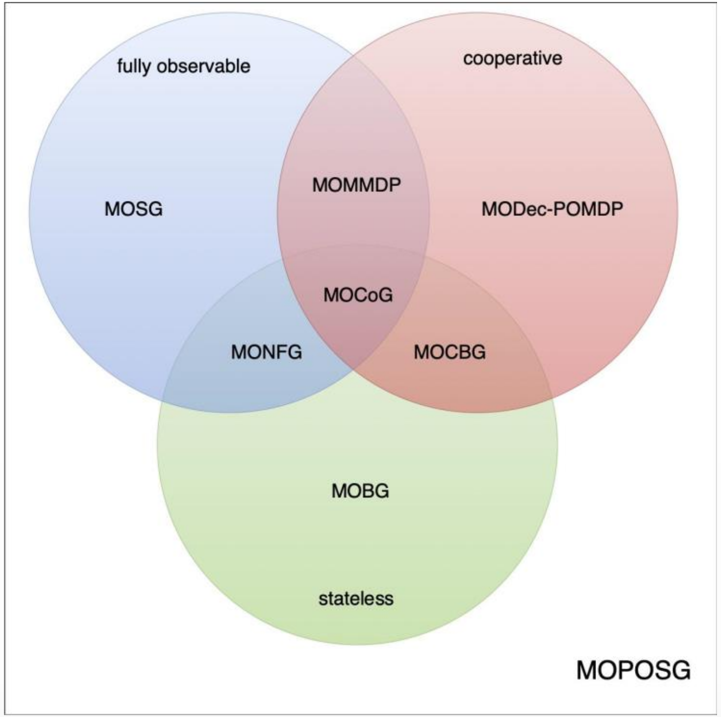
- User utility determines what is desirable for agents
- Stems from meaningful objectives (to the user)
  - Explainable
  - E.g., euros, minutes
- Identifying objectives
  - And then events that trigger rewards
- Decision-theoretic problem setting

MOPOSG



Models:
On the basis of rewards (in objectives) and observations (about states).

MOPOSG



Models:
On the basis of rewards (in objectives) and observations (about states).

But utility is not yet modelled!

Life is still not simple

- What are your objectives for your current research project?
  - Publishing asap?
  - Quality of conference/journal?
  - Collaboration potential?
  - Flag-posting?
  - Increasing funding potential?
  - Finishing your PhD?

- Setting?

Life is still not simple at all?

- What are your objectives for your current research project?
  - Publishing asap?
  - Quality of conference/journal?
  - Collaboration potential?
  - Flag-posting?
  - Increasing funding potential?
  - Finishing your PhD?

- Truly cooperative though?

Utility-based approach

- Utility function, u_i, maps vector to scalar utility
- Total preference order (can always make a decision between alternatives)
- Utility determines what is optimal within available policies

Utility-based approach

- Solution should be derived from utility
  - Not axiomatically assumed

- This leads to a taxonomy based on rewards and utilities (Part 2)

How to deal with MO problems

- Collect available information about user utility.
- Decide which policies (e.g., stochastic vs deterministic) are allowed.
- Derive the optimal solution concept from the resulting information of the first two points.
- Select or design an algorithm that fits the solution concept.
- When multiple policies are required for the solution, design a method for the user to select the desired policy among these optimal policies.

Part 2 - Structuring the MOMADM field

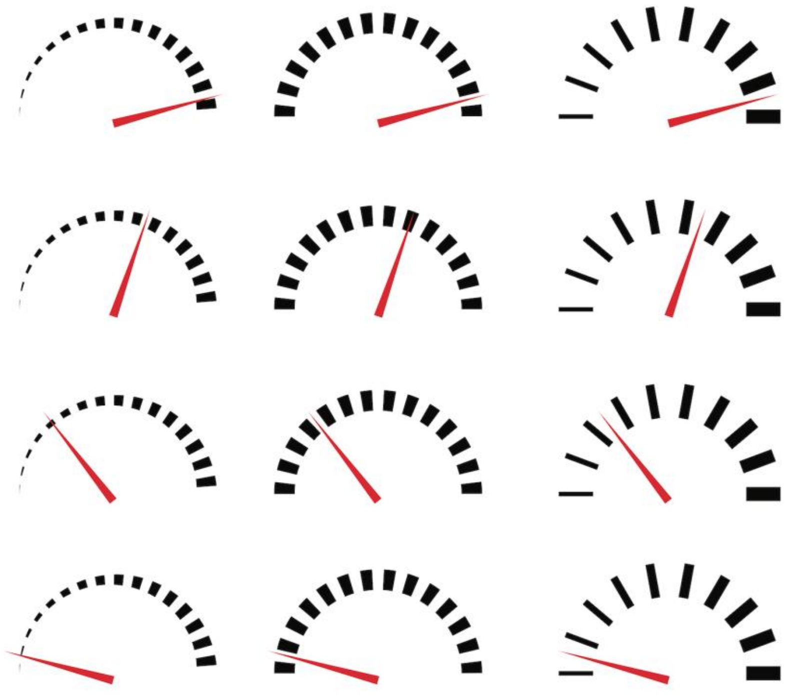
Taxonomy and solution concepts

Optimisation criteria

- Vectorial reward function

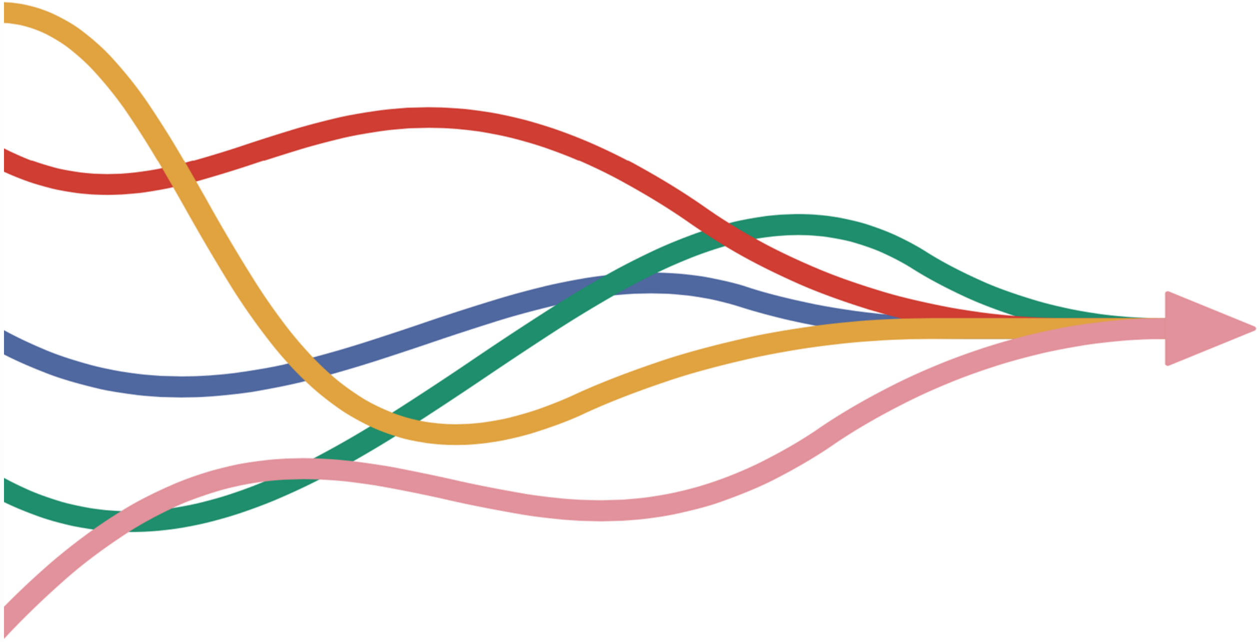
- Utility-based perspective

$$u_i : \mathbb{R}^d \rightarrow \mathbb{R}$$



Optimisation criteria

Optimisation criteria



Optimisation criteria

- Expected Scalarised Returns (ESR)
  - Calculate the expectation of the utility from the payoffs
  - Utility of an individual policy execution

Optimisation criteria

- Expected Scalarised Returns (ESR)
  - Calculate the expectation of the utility from the payoffs
  - Utility of an individual policy execution

- Scalarised Expected Returns (SER)
  - Calculate the utility of the expected payoff
  - Utility of the average payoff from several executions of the policy

Optimisation criteria

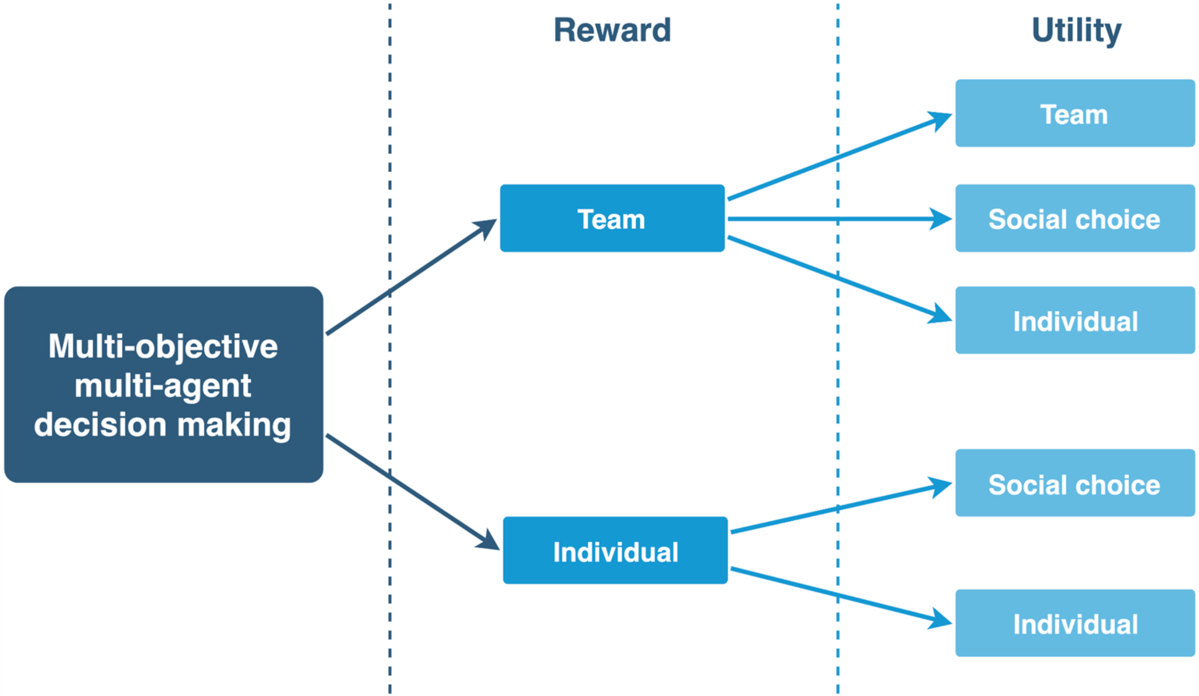
- Expected Scalarised Returns (ESR)

$$V_u^\pi = \mathbb{E}\left[u\left(\sum_{t=0}^{\infty} \gamma^t \mathbf{r}_t \right) \mid \pi, \mu_0 \right]$$

- Scalarised Expected Returns (SER)

$$V_u^\pi = u\left(\mathbb{E}\left[\sum_{t=0}^{\infty} \gamma^t \mathbf{r}_t \mid \pi, \mu_0 \right] \right)$$

Taxonomy



Rădulescu, R., Mannion, P., Roijers, D. M., & Nowé, A. (2020). Multi-objective multi-agent decision making: a utility-based analysis and survey. *Autonomous Agents and Multi-Agent Systems*, *34*(1), 1-52.

Taxonomy

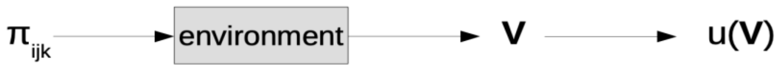

Reward
Utility
Multi-objective multi-agent decision making
Team
Individual
Team
Social choice
Individual
Social choice
Individual
ESR / SER

Examples - Team Reward

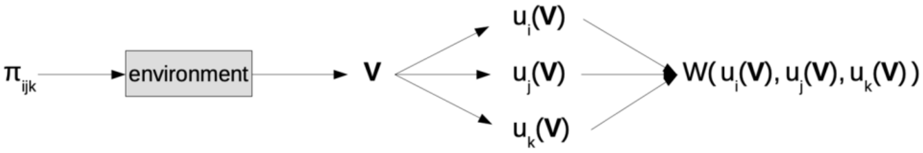
- Team utility
  - a company that aims to be environmentally responsible, while maximising profits

Examples - Team Reward

- Team utility
  - a company that aims to be environmentally responsible, while maximising profits



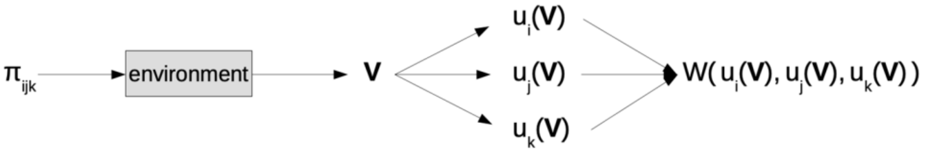
- Social Choice
  - highway tolls to regulate traffic



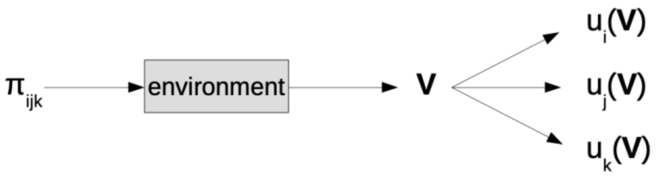
Examples - Team Reward

- Team utility
  - a company that aims to be environmentally responsible, while maximising profits

- Social Choice
  - highway tolls to regulate traffic

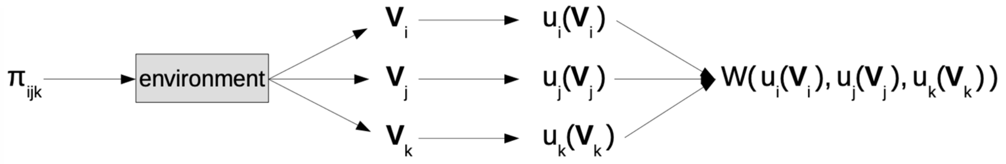


- Individual utility
  - participating in an event/planning a holiday together with your friends



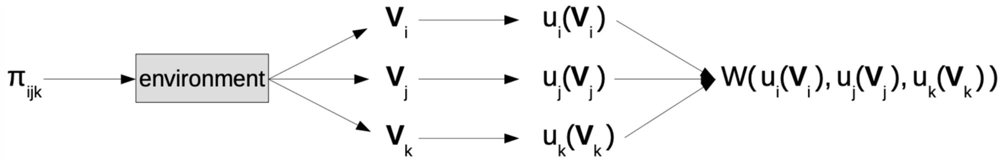
Examples - Individual Reward

- Social choice
  - bidding fee auctions

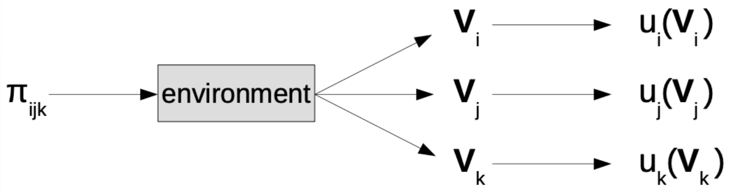


Examples - Individual Reward

- Social choice
  - bidding fee auctions



- Individual utility
  - participating in city traffic, work commutes



Solution concepts

		Utility: Team	Utility: Social choice	Utility: Individual
Reward	Team	Coverage sets	Mechanism design	Coverage sets (+ Negotiation) Equilibria and stability concepts
Reward	Individual		Mechanism design	Equilibria and stability concepts Coverage Sets as best responses

Coverage sets

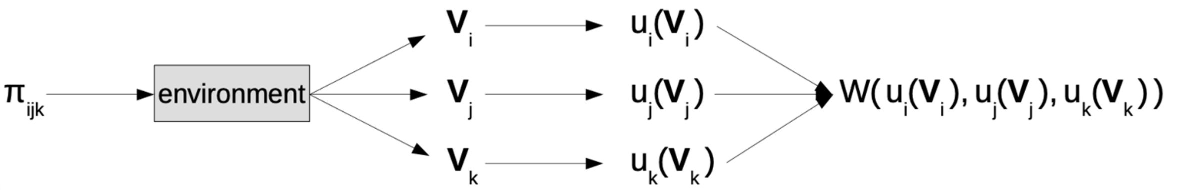
- Contain at least one optimal policy for each possible utility function
- **TRTU**: rewards and derived utility is shared between agents, with one utility function selected during execution
- **TRIU**: agent can (contractually) agree which policy to execute
- **IRIU**: set of possible best responses to the behaviour of other agents

		Utility		
		Team	Social choice	Individual
Reward	Team	Coverage sets	Mechanism design	Coverage sets (+ Negotiation) Equilibria and stability concepts
	Individual		Mechanism design	Equilibria and stability concepts Coverage Sets as best responses

Social Welfare and Mechanism Design

		Utility		
		Team	Social choice	Individual
Reward	Team	Coverage sets	Mechanism design	Coverage sets (+ Negotiation) Equilibria and stability concepts
	Individual		Mechanism design	Equilibria and stability concepts Coverage Sets as best responses

- System perspective: what is a socially desirable outcome



Design a system that forces agents to the truthful about their utilities and leads to optimal solution under W

Equilibria and stability concepts

		Utility		
		Team	Social choice	Individual
Reward	Team	Coverage sets	Mechanism design	Coverage sets (+ Negotiation) Equilibria and stability concepts
	Individual		Mechanism design	Equilibria and stability concepts Coverage Sets as best responses

- Stable outcomes from which self-interested agents have no incentive to deviate
- Nash equilibria, correlated equilibria, cyclic equilibria, coalition formation

Nash Equilibrium

- No agent can improve their utility by unilaterally deviating from the joint strategy π^{NE}

- Nash equilibrium under SER:

$$\mathbb{E}u_i\left[\mathbf{p}_i(\pi_i^{NE}, \pi_{-i}^{NE})\right] \geq \mathbb{E}u_i\left[\mathbf{p}_i(\pi_i, \pi_{-i}^{NE})\right]$$

- Nash equilibrium under ESR:

$$u_i\left[\mathbb{E}\mathbf{p}_i(\pi_i^{NE}, \pi_{-i}^{NE})\right] \geq u_i\left[\mathbb{E}\mathbf{p}_i(\pi_i, \pi_{-i}^{NE})\right]$$

Other solution concepts

- Cyclic Nash equilibria
    - No agent can improve their utility by unilaterally deviating from the joint cyclic strategy

- Correlated equilibria
    - Correlated strategy - probability vector σ on $\mathcal{A}$
    - External mechanism
    - No agent can improve their utility by unilaterally deviating from the recommendation of the correlated signal

Example

$$u(p_1, p_2) = p_1 \cdot p_2$$

	A	B
A	(10, 2); (10, 2)	(0, 0); (0, 0)
B	(0, 0); (0, 0)	(2, 10); (2, 10)

Example - Nash equilibrium

$$u(p_1, p_2) = p_1 \cdot p_2$$

$$u(10, 2) = 10 \cdot 2 = 20$$

	A	B
A	(10, 2); (10, 2)	(0, 0); (0, 0)
B	(0, 0); (0, 0)	(2, 10); (2, 10)

Example - Cyclic Nash equilibrium

$$u(p_1, p_2) = p_1 \cdot p_2$$

	A	B
A	(10, 2); (10, 2)	(0, 0); (0, 0)
B	(0, 0); (0, 0)	(2, 10); (2, 10)

$$u(6, 6) = 6 \cdot 6 = 36$$

- Joint cyclic strategy
  - Player 1: {A, B}
  - Player 2: {A, B}

Example - Correlated equilibrium

$$u(p_1, p_2) = p_1 \cdot p_2$$

	A	B
A	(10, 2); (10, 2)	(0, 0); (0, 0)
B	(0, 0); (0, 0)	(2, 10); (2, 10)

$$u(6,6) = 6 \cdot 6 = 36$$

- Correlated strategy σ
  - 50% (A, A)
  - 50% (B, B)

Part 3 - SOTA

Latest results and open challenges

Multi-Objective Normal Form Games

- Introduced by Blackwell in 1956
- MONFG - tuple (N , A, **p**), with $n \geq 2$ and $C \geq 2$ objectives, where:
  - $N = \{1, ..., n\}$ – set of players
  - $A = A_1 \times \dots \times A_n$ – set of actions
  - $\mathbf{p} = (\mathbf{p}_1, ..., \mathbf{p}_n)$ – vectorial payoffs

(Im)balancing Act Game

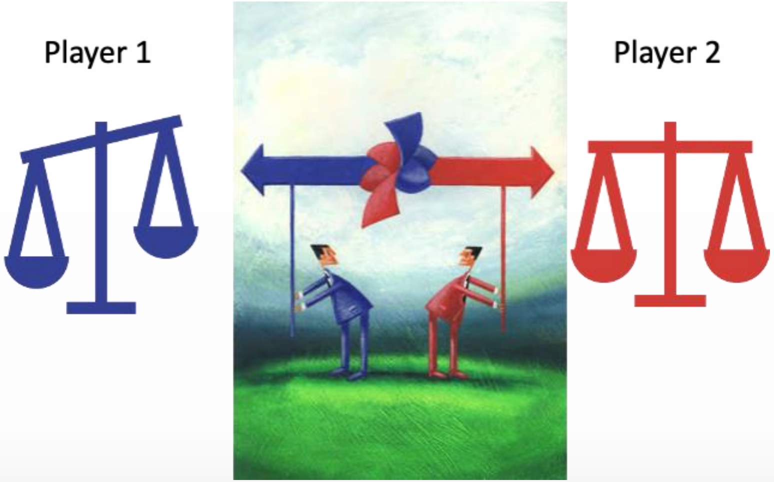
- 2 players, 2 objective
- Same payoff vector for both players

$$u_1([p_1, p_2]) = p_1^2 + p_2^2$$
$$u_2([p_1, p_2]) = p_1 \cdot p_2$$

	L	M	R
L	[4,0]	[3,1]	[2,2]
M	[3,1]	[2,2]	[1,3]
R	[2,2]	[1,3]	[0,4]

Theoretical considerations

- In finite MONFGs, where each agent seeks to maximise the utility under **SER**, **Nash equilibria need not exist.**



	L	M	R
L	[4,0]	[3,1]	[2,2]
M	[3,1]	[2,2]	[1,3]
R	[2,2]	[1,3]	[0,4]

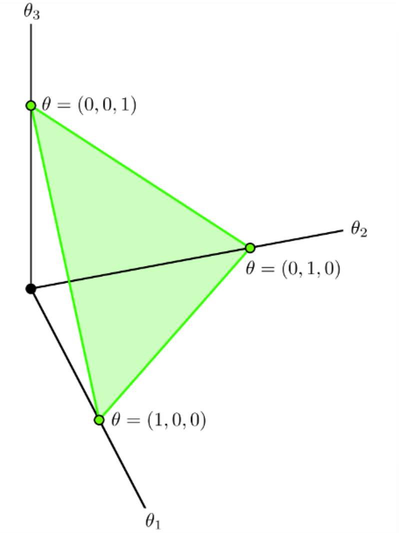
$$u_1([p_1, p_2]) = p_1^2 + p_2^2$$
$$u_2([p_1, p_2]) = p_1 \cdot p_2$$

Rădulescu, R., Mannion, P., Zhang, Y., Roijers, D. M., & Nowé, A. (2020). A utility-based analysis of equilibria in multi-objective normal-form games. *The Knowledge Engineering Review*, *35*.

Novel intuition

- Every MONFG with continuous utility functions can be reduced to a continuous game

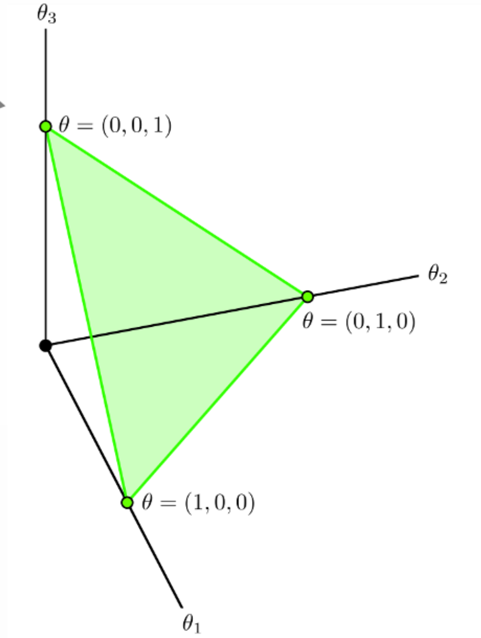
- Continuous games:
  - Single objective
  - Infinite number of pure strategies
  - Reuse utility functions



Röpke, W., Roijers, D. M., Nowé, A., & Rădulescu, R. (2021). On Nash Equilibria in Normal-Form Games With Vectorial Payoffs. *arXiv preprint arXiv:2112.06500*.

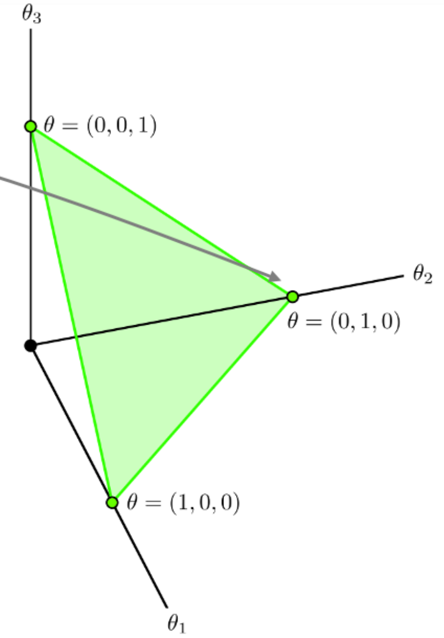
Novel intuition

	L	M	R
L	[4,0]	[3,1]	[2,2]
M	[3,1]	[2,2]	[1,3]
R	[2,2]	[1,3]	[0,4]



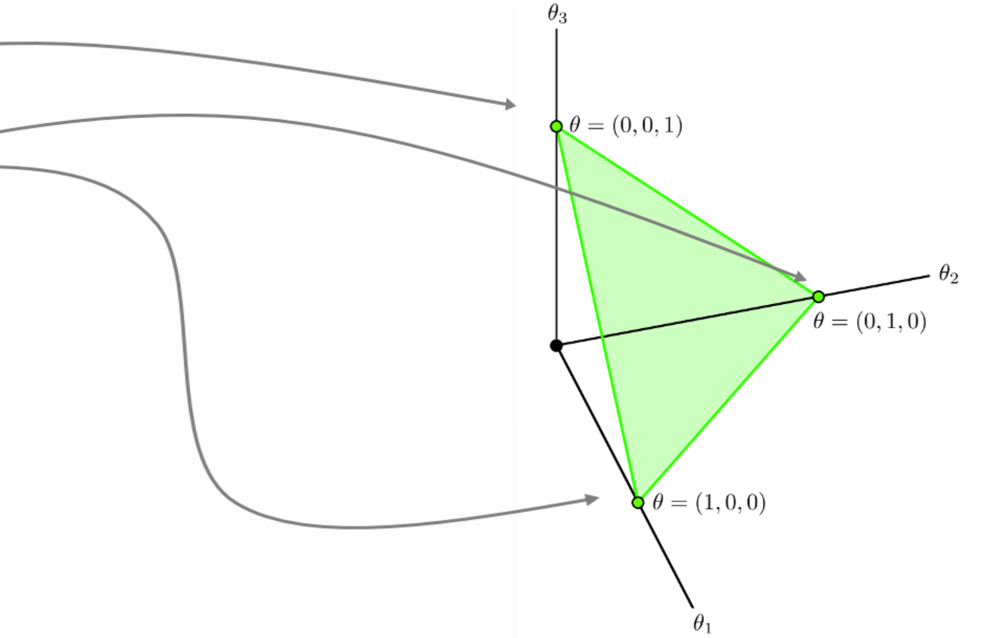
Novel intuition

	L	M	R
L	[4,0]	[3,1]	[2,2]
M	[3,1]	[2,2]	[1,3]
R	[2,2]	[1,3]	[0,4]



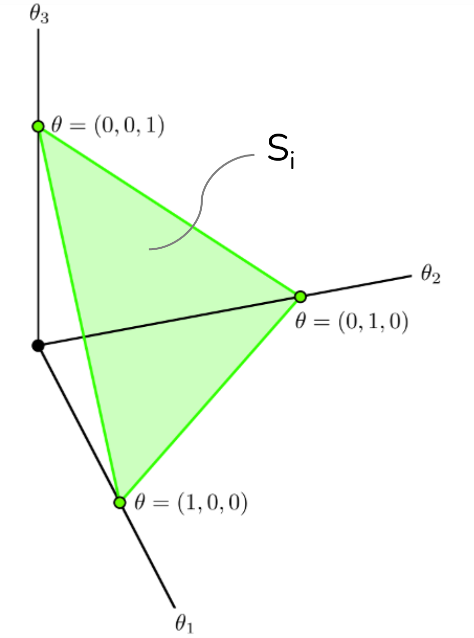
Novel intuition

	L	M	R
L	[4,0]	[3,1]	[2,2]
M	[3,1]	[2,2]	[1,3]
R	[2,2]	[1,3]	[0,4]



Novel intuition

	L	M	R
L	[4,0]	[3,1]	[2,2]
M	[3,1]	[2,2]	[1,3]
R	[2,2]	[1,3]	[0,4]

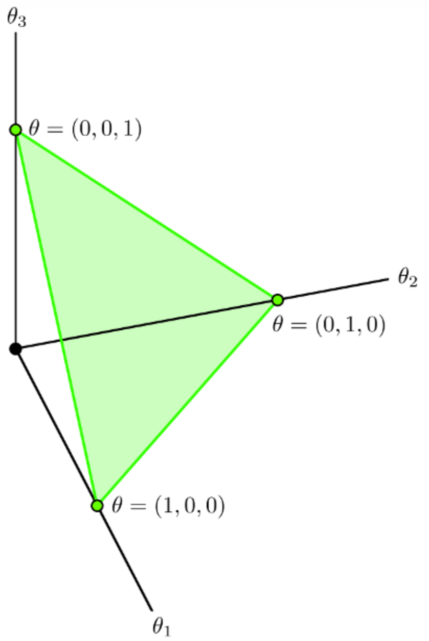


Novel intuition

- Mixed strategy equilibria in the MONFG are pure strategy equilibria in the continuous game

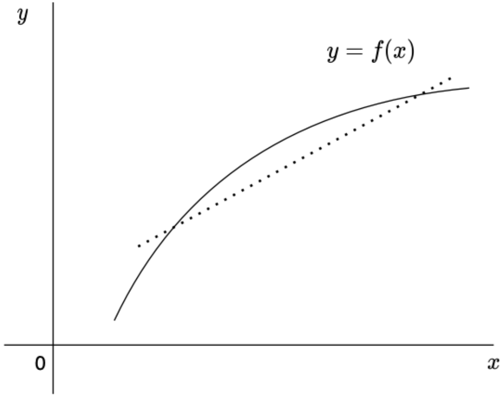
- Continuous games are not guaranteed to have a pure strategy Nash equilibrium

  ▶ Nash equilibria are not guaranteed in MONFGs



NE Existence Guarantees

- Existence is guaranteed with (quasi)concave utility functions
  - Used in economics as well
  - Represents “well-behaved” preferences
- Intuition
  - MONFGs can be reduced to continuous games
  - In these game it is known that a pure strategy NE exists when assuming only quasiconcave utility functions
  - This equilibrium is also an equilibrium in the original MONFG



Röpke, W., Roijers, D. M., Nowé, A., & Rădulescu, R. (2021). On Nash Equilibria in Normal-Form Games With Vectorial Payoffs. *arXiv preprint arXiv:2112.06500*.

Non-existence

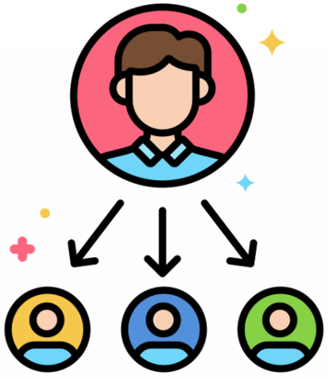
- We can show that no Nash equilibrium exists in this game
  - With **strict convex utility functions**

	A	B
A	(2, 0); (1, 0)	(1, 0); (0, 2)
B	(0, 1); (2, 0)	(0, 2); (0, 1)

$$u_1(p_1, p_2) = u_2(p_1, p_2) = p_1^2 + p_2^2$$

Commitment and Cyclic Strategies

- Commitment
  - One or more players commit to playing a specific strategy
  - Other players condition their own strategies on this commitment

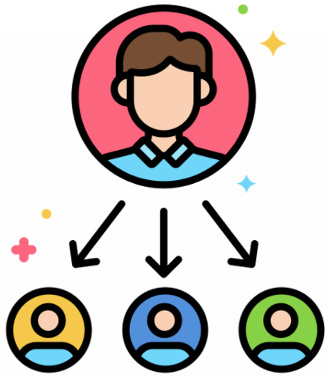


- Leadership equilibria (in two-player games)
  - The leader cannot improve their utility given that the follower plays a best-response
- Weak/strong leadership equilibria
  - Prescribes how an opponent selects their best-response

Röpke, W., Roijers, D. M., Nowé, A., & Rădulescu, R. (2021). Preference Communication in Multi-Objective Normal-Form Games. *Neural Computing and Applications (in press).*

Theoretical considerations

- Commitment can be strictly better than all Nash equilibria
  - Commit may avoid the “fixed-point death trap”



$$u(p_1, p_2) = p_1 \cdot p_2$$

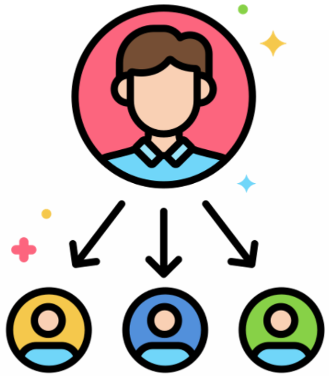
	A	B
A	(10, 2); (10, 2)	(0, 0); (0, 0)
B	(0, 0); (0, 0)	(2, 10); (2, 10)

Nash equilibrium

$$u(10, 2) = 10 \cdot 2 = 20$$

Theoretical considerations

- Commitment can be strictly better than all Nash equilibria
  - Commit may avoid the “fixed-point death trap”



The optimal mix is to play 50% (A, A) and 50% (B, B)

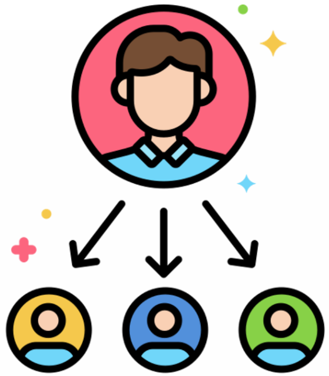
$$u_1\left(\frac{10+\ 2}{2},\frac{2+10}{2}\right)=u_1(6,6)=36$$

$$u(p_1,p_2)=p_1\cdot p_2$$

	A	B
A	(10, 2); (10, 2)	(0, 0); (0, 0)
B	(0, 0); (0, 0)	(2, 10); (2, 10)

Theoretical considerations

- Commitment can be strictly better than all Nash equilibria
  - Commit may avoid the “fixed-point death trap”



The optimal mix is to play
50% (A, A) and 50% (B, B)

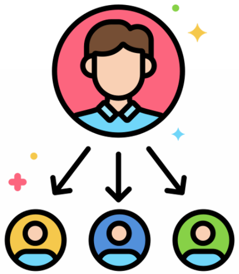
$$u_1\left(\frac{10+\ 2}{2},\frac{2+10}{2}\right) = u_1(6,6) = 36$$

- Joint cyclic strategy
  - Player 1: {A, B}
  - Player 2: {A, B}

$$u(p_1,p_2) = p_1 \cdot p_2$$

	A	B
A	(10, 2); (10, 2)	(0, 0); (0, 0)
B	(0, 0); (0, 0)	(2, 10); (2, 10)

Theoretical considerations



- Commitment is not guaranteed to be as good as a Nash equilibrium
  - If a player commits to a strategy, a malicious player might exploit this
  - This has implications for a range of real-world applications
- Cyclic Nash equilibria may exist when no stationary equilibrium exists
  - Stable solutions can still exist
  - Provides a valid alternative for the goal of a learning algorithm

Relations between optimisation criteria

- **Mixed strategies**
  - **No relation** between both optimisation criteria **in general**

	A	B
A	(1, 0); (1, 0)	(0, 1); (0, 1)
B	(0, 1); (0, 1)	(-10, 0); (-10, 0)

Multi-objective reward vectors

	A	B
A	0.1; 0.1	0; 0
B	0; 0	-0.1; -0.1

Scalarised utility for both agents

No sharing of number of equilibria or equilibria themselves

Relations between optimisation criteria

- **Pure strategies**
  - Pure strategy equilibrium under SER is also one under ESR
  - Bidirectional when assuming (quasi)convex utility functions

- We can extend this to **blended settings**
  - Pure strategy equilibrium under SER is also one in any blended setting
  - Bidirectional when assuming (quasi)convex utility functions

Open questions

- Commitment and cyclic strategies
  - When can we guarantee that commitment cannot be exploited?
  - What is the link between correlated equilibria and hierarchical equilibria?
  - How to extend the Stackelberg game model to n-player games?
  - Open computational problems
    - Algorithm for learning or computing optimal commitment strategies?
    - How to learn hierarchical strategies?

Open questions

- Results for more complex (e.g., sequential, partially observable) settings
- Integrated pipelines for planning -> negotiation -> execution
- Utility modelling
- Strategic disclosure of utility information to the other agents
- Benchmarks

Thank you for listening

- Feel free to ask any questions now
- Or drop us a message

This tutorial was based (primarily) on

- Rădulescu, R., Mannion, P., Roijers, D. M., & Nowé, A. (2020). Multi-objective multi-agent decision making: a utility-based analysis and survey. Autonomous Agents and Multi-Agent Systems, 34(1), 1-52.
- Rădulescu, R., Mannion, P., Zhang, Y., Roijers, D. M., & Nowé, A. (2020). A utility-based analysis of equilibria in multi-objective normal-form games. The Knowledge Engineering Review, 35.
- Rădulescu, R. (2021). Decision Making in Multi-Objective Multi-Agent Systems: A Utility-Based Perspective. Brussels: Crazy Copy Center Productions.
- Röpke, W., Roijers, D. M., Nowé, A., & Rădulescu, R. (2021). Preference Communication in Multi-Objective Normal-Form Games. Neural Computing and Applications (in press).
- Röpke, W., Roijers, D. M., Nowé, A., & Rădulescu, R. (2021). On Nash Equilibria in Normal-Form Games With Vectorial Payoffs. arXiv preprint arXiv:2112.06500.