Game Theory for Morality

Amir Ahmad Habibi | April 2019

How can
Game Theory
be used in
Moral Philosophy?

What we are going to discuss

[0] Some History

[1] Functionalism
[2] Contractarianism
[3] Evolutionary GT

Some History

Some History

Richard B. Braithwaite

"Theory of Games as a Tool
for the Moral Philosopher"
1955

Some History

"Theory of Games and Economic Behaviour"
1944

Von Neumann & Morgenstern

"The bargaining problem"
1950

John F. Nash

"Theory of Games as a Tool for the Moral Philosopher"
1955

R. B. Braithwaite

Some History

Cool GT Timeline:
http://euler.fd.cvut.cz/pred
mety/teorie_her/histf.html

Some History

Thomas
Hobbes
1588 - 1679

David
Hume
1711 - 1776

Some History

Richard B. Braithwaite

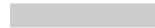His prediction hasn't yet happened :(

Some History

Bruno Verbeek 2002

[1] Functionalism

[2] Contractarianism

[3] Evolutionary GT

1st approach
Functionalism

1. Functionalism

Edna Ullmann-Margalit

1977

The Emergence of Norms

moral norms enable agents to
cooperate and coordinate their actions
in situations where
the pursuit of self-interest prevents this.

1. Functionalism

Edna Ullmann-Margalit
The Emergence of Norms 1977

Artillerist #2

		stay	flee
Artillerist #1	stay	2 2	3 0
	flee	0 3	1 1

1. Functionalism

Edna Ullmann-Margalit
The Emergence of Norms 1977

The solution?
 - **Chain** each other to the guns

> **Morality** commits agents
> to avoid Pareto-inefficient
> or sub-optimal outcomes.

1. Functionalism
The Problems:

- It does not explain the emergence and persistence of morality

- We don't know if morality coincides with the Pareto-superior outcomes.

- It assumes the conflict of morality with individual rationality

Bruno Verbeek 2002

1. Functionalism

<

2. Contractarianism

3. Evolutionary GT

Bruno Verbeek 2002

2nd approach
Contractarianism
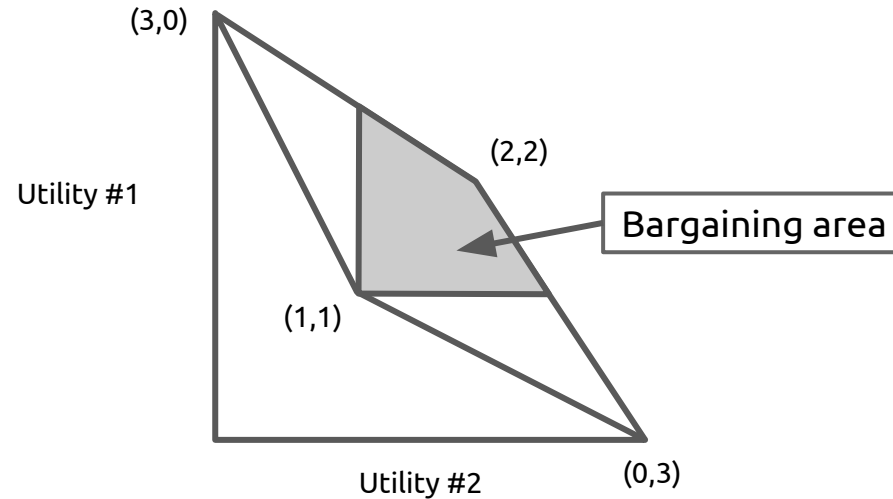
2. Contractarianism

Morality as a Bargaining Process:

Rational agents
in a suitably idealised
bargaining situation
will agree on a specific
distribution of benefits
of cooperation.

2. Contractarianism

The Artillerists Bargaining

2. Contractarianism

What would the distribution be?

- Cooperative game theory

- Non-cooperative game theory

——

2. Contractarianism

David Gauthier

Morals by Agreement 1986

Tries to answer "why be Moral?"

2. Contractarianism

David Gauthier

1986
Morals by Agreement

- Constrained maximisation

- Disposition to cooperate
  with other cooperators

3rd approach

Evolutionary-gametheorianism

3. EvoGTanism

Differences:

- Morality is a **side-effect** of interactions of agents

- It emerges from repeated interactions of **small groups**

- Not assuming **full information** and **full rationality**

3. EvoGTanism

Jean-Jacques Rousseau

1754

Discours sur l'origine et les fondements de l'inégalité parmi les hommes
A Discourse Upon the Origin and Foundation of the Inequality Among Mankind

> Stag Hunt <

3. EvoGTanism

Jean-Jacques Rousseau
Stag Hunt   **1754**

	Stag	Hare
Stag	3 3	2 0
Hare	0 2	2 2

#1

3. EvoGTanism

Jean-Jacques Rousseau
Stag Hunt   **1754**

Pareto-efficient
But unstable

#2

	Stag	Hare
Stag	3 3	2 0
Hare	0 2	2 2

#1

3. EvoGTanism

According to EvoGT:

Since most of our actual
norms are often stable,
they might be
Pareto-inefficient
unlike Contractarian norms.

3. EvoGTanism

Achievements:

- Explains the emergence & persistence of moral norms

- Has recovered many existing moral norms

- Provides an answer to "Why be moral?"

3. EvoGTanism

Drawback:

- Focus on explanation but not justifications

- Immoral norms!

- Moral skepticism

Epilogue

Moral theorists are against

This rather one-dimensional picture of rational man.

Kreps & **Wilson**

1982

Reputation and
Imperfect Information

Bertrand Russell

"It has been said that man is a rational animal. All my life I have been searching for evidence which could support this."

Future work

Synthesis of
the traditional Contractarian
and the recent EvoGT
approaches

Questions?