

Disobedience as a Mechanism of Change

David Burth Kurka¹, Jeremy Pitt¹, Peter Lewis², Alina Patelli² and Anikó Ekárt²

12th IEEE SASO, 3-7 Sep 2018, Trento Italy

¹Department of Electrical and Electronic Engineering, Imperial College London

²Aston Lab for Intelligent Collectives Engineering, Aston University

Imperial College
London


Aston University
Birmingham

Introduction

*The social concept of **disobedience** can be used as a mechanism for **adaptive and exploratory learning** in norm-governed multi-agent systems.*

Disobedience in Society

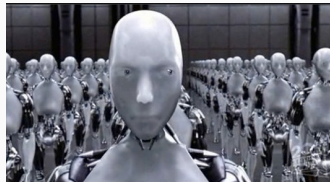
- **Disobedience** has been used historically as a form of resistance and call for change in unfair regimes.
- Among the factors that motivate widespread disobedience are:
 - **The iron law of oligarchy** - the tendency for a small group to emerge and run complex organisations for its own benefit, rather than the collective interest
 - **Principled violation of policy** - law enforcement relaxation, with selective common-sense non-application of sanctions



Figure 1: Suffragette's in UK

Disobedience as Learning Mechanism

- Can these notions be applied to **norm-governed MAS**?
- How to deal with systems with **disfunctional norms**?
- Can disobedience be a form of **collective learning**, enabling **adaptive systems**?

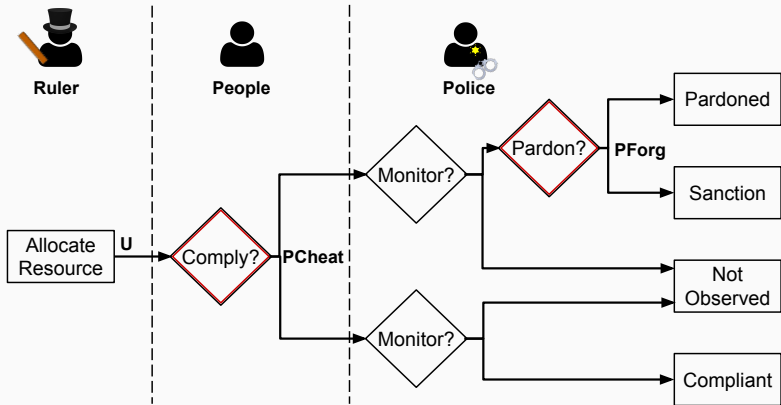


System Model



- Common-Pool Resource Management game
- Players and its roles:
 - **Rulers** - Resource allocation policy
 - **People** - Appropriation of resource
 - **Police** - Check if appropriation = allocation

Disobedience Levels



Learnable Variables

Rulers:

U - Corruption of allocation policy

$$U = f(time)$$

Police:

P_{forg} - Probability of sanction pardon as function of perceived fairness

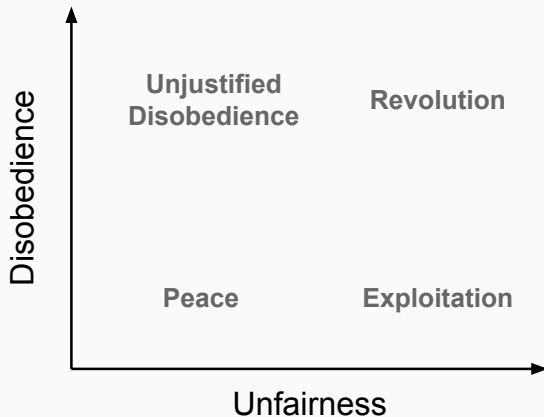
$$P_{forg} = f(fairness) \quad fairness \sim Gini(Satisf)$$

People:

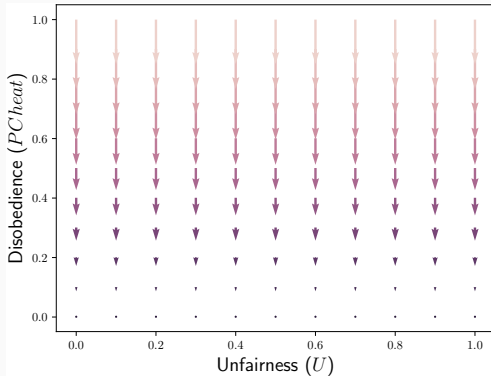
P_{Cheat} - Decision of compliance as reaction of system's feedback

$$P_{Cheat} = f(fairness_i) \quad \left\{ \begin{array}{l} \uparrow \text{ if sanctioned } \wedge \text{ unfair} \\ \downarrow \text{ if sanctioned } \wedge \text{ fair} \\ = \text{ if forgiven} \\ \uparrow \text{ if not caught} \\ \downarrow \text{ if compliant} \end{array} \right.$$

Experimental Results

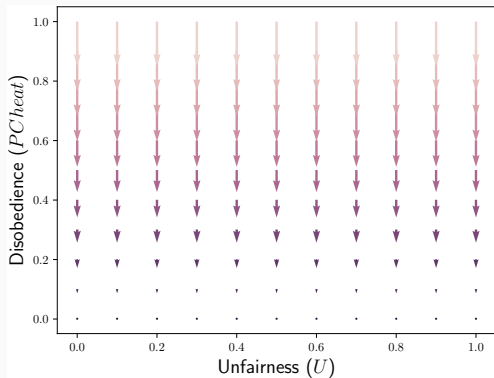


Base Game

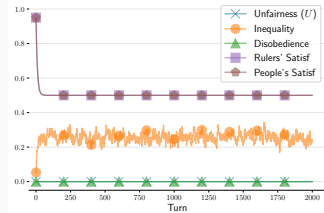


Without pardoning or reformation, People stay at the mercy of rulers, not having another rational choice than to obey, no matter how unfair is the current policy.

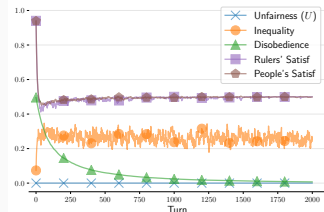
Base Game



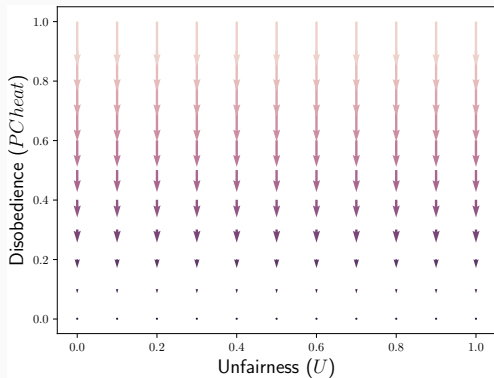
Peace:



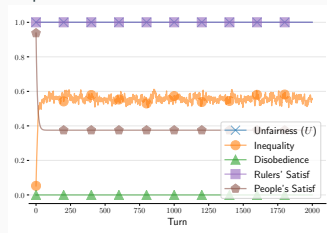
Unjustified Disobedience:



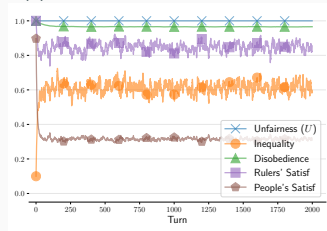
Base Game



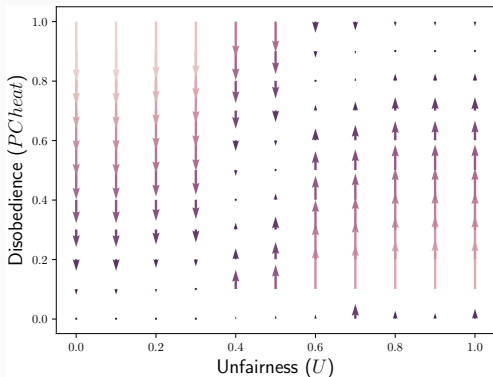
Exploitation:



Oppression:



First Extension - Pardoning

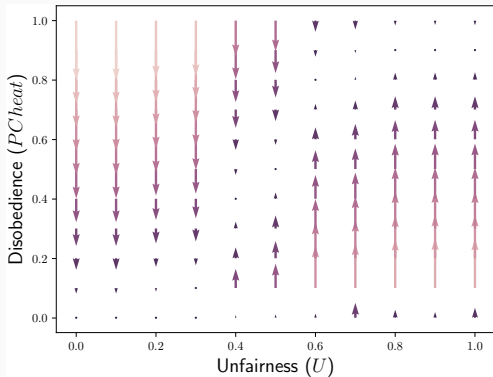


Obedience is conditioned to unfairness level:

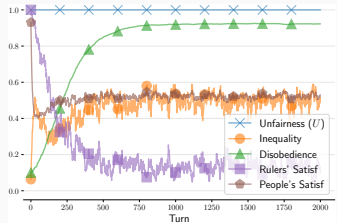
If fair, agents obey; if unfair, tendency to non-compliance

Police validates justified norms disobedience, pardoning transgressions

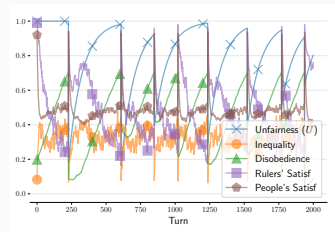
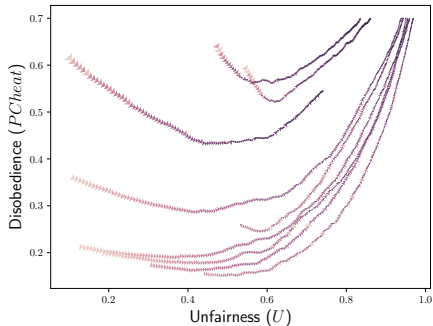
First Extension - Pardoning



Revolution:

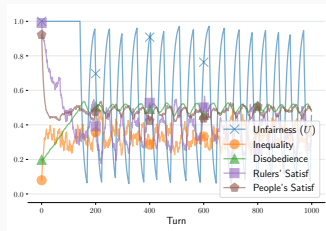
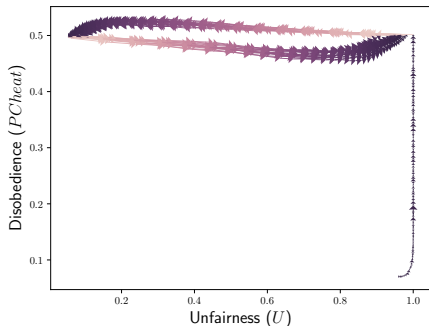


Second Extension - Reformation - Reformation Cycles



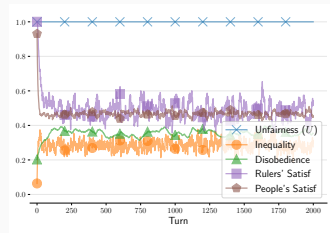
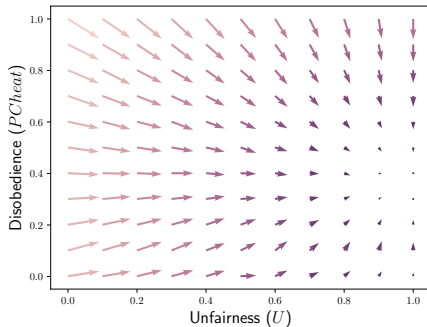
- Perpetual cycle of power change
- Alternation of power and higher overall satisfaction.
- Points out the need for post-revolution reconstruction.

Second Extension - Reformation - Popular Control



Knowing the Reformation Threshold \mathcal{T} , unfairness is manipulated so Rulers please People with just enough resources so their disobedience remains in controlled levels.

Second Extension - Reformation - Pragmatic Revolution



People disobey just enough to guarantee that satisfaction (and resource) is the same for all agents.

Conclusion

- **Disobedience** as a tool for adaptive learning in norm-governed MAS
 - Usual systems try to guarantee **obedience**, but not **governance fairness**
 - **pardoning mechanisms** justify **disobedience** as a revolutionary behaviour for justice restoration
 - Formal **reformation mechanisms** give power to non-Rulers and enables **cycles of change**
- Open questions / Future work:
 - Can subjective terms as *fairness, obedience, unjustified, revolution* be defined in different ways?
 - What are the effects of limited or subjective knowledge of events? How interpersonal factors such as *social influence, trust* and *misinformation* could affect the system?
 - How to escape from simple system reformation and move towards **post-revolution reconstruction**?

- National Council for Scientific and Technological Development (CNPq), Brazil
- Diverse collaborators



Thank you!

Questions?

- **Base Premises**
 - P1 - A norm is in place
 - P2 - Compliance is an individual decision
 - P3 - Existence of monitoring and sanctioning mechanisms
- **Disobedience Premises**
 - P4 - Fairness decays over time (corruption)
 - P5 - Possibility of pardon by law enforcers
- **Transformation Premises**
 - P6 - Prescribed ruler destitution and new government formation