

# Game Theory & Cooperation

## Course Overview

Human agents can be rational and human agents can be cooperative. They conform to neither description perfectly, but close enough that questions arise how these two features go together. When do our rationality and our cooperative nature support each other and when are they in tension? We look to game theory for an answer.

This course serves both the purpose of introducing the basics of game theory for interdisciplinary research and of exploring the dynamics of cooperation. Some games, such as the *Prisoners' Dilemma* or the *Stag Hunt*, have become famous for the problems they pose for cooperation. These well-known games, however, barely scratch the surface of what game theory can reveal about human cooperation, and cooperation in general. Game theory sheds light on conventions, social norms, and the evolution of cooperation. More radical contributions to game theory propose preference change and team reasoning as mechanisms for cooperation.

The course will be highly interdisciplinary. It draws on work in philosophy of language, the social sciences, biology, social ontology, and more.

## Intended Audience

Advanced undergraduate students, early post-graduate students. Although of help, prior knowledge of game theory or decision theory is not required.

## Course Aims

The course material provides the participants with the following:

- An understanding of basic game theory, including important games such as the *Prisoners' Dilemma* and the *Stag Hunt* and equilibria concepts.
- Familiarity with various game theoretical analyses of social norms.
- An introduction to evolutionary game theory and its approach to cooperation.
- A skillset for conducting interdisciplinary work, ranging from economics and the social sciences to biology and practical philosophy.
- Awareness of how philosophy and game theory can provide an account of our everyday cooperation.

## Assessment

Students are given the freedom to choose between two modes of assessment: A short essay (3000-4000 words) plus a two-hour exam, or a more extensive essay (5000-6000 words).

## Online Resources

Game theory is a hot topic with many helpful resources available online for free. If you want to refresh your memory or dig deeper, consider the following:

Introductory Game Theory Videos

<https://www.youtube.com/playlist?list=PLqekkRyYeow3cR9U4c4wkIekm2pXxORPn>

The Evolution of Trust

<http://ncase.me/trust/>

Open Yale Course on Game Theory

<http://oyc.yale.edu/economics/econ-159>

David K. Levine's Site on Game Theory

<http://dklevine.com/>

## **Week 1 – Introduction to Game Theory I**

### *Lecture 1*

In the first lecture, we discuss the practicalities of the course and elaborate our research questions:

1. How do we maintain cooperative social norms?
2. Can game theory tell us about the evolution of cooperation?
3. Do our motivations change to support cooperation?
4. Can game theory capture the human capacity to act together as a team?

### *Lecture 2*

We investigate decision theory as the fundament of game theory. The assumptions built into the axioms of decision theory carry over to game theoretic models of interaction, thus we better have a good look at them and see whether we find them plausible. We also start to become familiar with the formal conception of action used in game theory.

### **Mandatory Reading**

Peterson, M. (2009). *An Introduction to Decision Theory*. Cambridge University Press. Chapter 2. pp. 17-39

Tadelis, S. (2013). *Game Theory. An Introduction*. Princeton University Press. pp. 3-40.

### **Suggested Reading**

Bradley, R. (2014). *Decision Theory: A Formal Philosophical Introduction*. Online:

[http://personal.lse.ac.uk/bradleyr/pdf/Handbook%20-%20Decision%20theory%20\(revised\).pdf](http://personal.lse.ac.uk/bradleyr/pdf/Handbook%20-%20Decision%20theory%20(revised).pdf)

Steele, K., & Stefánsson, H. O. (2016). Decision Theory. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University. Retrieved from

<https://plato.stanford.edu/archives/win2016/entries/decision-theory/>

## Week 2 – Introduction to Game Theory II

### *Lecture 3*

Game theory applies the formal methods of decision theory to social interaction. In this lecture we take the first steps to master game theoretic models. We learn the tree and matrix representation of games and look at the famous *Prisoner's Dilemma*, which allows us to introduce the concept of a dominant strategy.

### *Lecture 4*

We delve deeper into game theory, consider further games such as the *Stag Hunt* and expand our toolbox with the concept of equilibrium states. This lecture will also serve to recap the formal models we have learned and make sure we all understand the basics going forward.

### **Mandatory Reading**

Reading: Peterson, M. (2009). *An Introduction to Decision Theory*. Cambridge University Press. Chapters 11 & 12. pp. 212-262

### **Suggested Reading**

Binmore, K. G. (2007). *Game Theory: A Very Short Introduction*. Oxford University Press.

Peterson, M (2015), *The Prisoner's Dilemma*, Cambridge University Press.

Ross, D. (2016). Game Theory. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University. Retrieved from:  
<https://plato.stanford.edu/archives/win2016/entries/game-theory/>

Tadelis, S. (2013). *Game Theory. An Introduction*. Princeton University Press. pp.79-126.

## Week 3 – Conventions. A First Look

### Lecture 5

David Lewis' 1969 *Convention* was a major break-through in modelling human cooperation using the means of game theory. His original aim was to establish that linguistic meaning can be conventional, although it did not literally arise by people convening. The impact of this book, however, went far beyond Lewis' original goal. We also encounter Lewis discussion of coordination games.

### Lecture 6

In *Convention* David Lewis' introduced the notion of common knowledge. It proved a major contribution to game theory but has also influenced accounts of cooperation not drawing on game theory. We consider its impact on a variety of games and their equilibria.

### Mandatory Reading

Lewis, D. (2011). *Convention: A Philosophical Study*. Blackwell. Chapter 1. pp. 5-51.

Tadelis, S. (2013). *Game Theory. An Introduction*. Princeton University Press. Chapter 4. pp. 59-78.

### Suggested Reading

Cubitt, R. P., & Sugden, R. (2003). Common Knowledge, Salience and Convention: A Reconstruction of David Lewis' Game Theory. *Economics & Philosophy*, 19(2), 175–210.

Vanderschraaf, P. (1998). Knowledge, Equilibrium and Convention. *Erkenntnis*, 49(3), 337–369.  
<https://doi.org/10.1023/A:1005461514200>

Vanderschraaf, P., & Sillari, G. (2014). Common Knowledge. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*. Retrieved from

<https://plato.stanford.edu/archives/spr2014/entries/common-knowledge/>

## Week 4 – Conventions and Norms I

### *Lecture 7*

We look at Lewis' short discussion of social norms in his *Conventions* and consider criticisms raised in the literature. Does Lewis' proposal fail to capture the motivational force of social norms? Furthermore, Lewis's conventions only work for coordination games, but social norms might also apply to other games. In the next lecture, we find a theory taking that into account.

### *Lecture 8*

Ullmann-Margalit contributed one of the first discussions relating game theory and norms. In contrast to Lewis, she distinguished norms from conventions. Her account does not presuppose that only coordination games underlie social norms. We look at her proposal that some social norms serve to overcome the mutual defection in Prisoners' Dilemma.

### **Mandatory Reading**

Lewis, D. (2011). *Convention: A Philosophical Study*. Blackwell. pp. 97-100.

Ullmann-Margalit, E. (2015). *The Emergence of Norms*. Oxford University Press. Chapter 2. pp. 18-73.

### **Suggested Reading**

Gilbert, M. (1989). *On Social Facts*. Princeton University Press. Chapter 6. pp. 315-407.

Ullmann-Margalit, E. (2015). *The Emergence of Norms*. Oxford University Press. Chapter 3. pp. 74-133.

## Week 5 – Conventions and Norm II

### Lecture 9

Christina Bicchieri's work stands in the tradition of Lewis and Ullmann-Margalit but has a much broader scope and goes into more details. She uses game theory to present her own analysis of social norms, at the centre of which stand conditional preferences and equilibria. We consider Bicchieri's typology of norms and conventions.

### Lecture 10

We wrap up our discussion of social norms, by considering real world cases. Bicchieri and others have applied their game theoretical frameworks to specific examples such as norms of corruption. Can we explain how corruption persists in societies and how it changes by modelling it using our game theoretic tools?

### Mandatory Reading

Bicchieri, C. (2005). *The Grammar of Society: The Nature and Dynamics of Social Norms*. Cambridge University Press. Chapter 1. pp. 1-54.

### Suggested Reading

Bicchieri, C. (2017). *Norms in the Wild: How to Diagnose, Measure, and Change Social Norms*. Oxford University Press.

Bicchieri, C., & Funcke, A. (2018). Norm Change: Trendsetters and Social Structure. *Social Research: An International Quarterly*, 85(1), 1–21.

Bicchieri, C., Jeffrey, R., & Skyrms, B. (2009). *The Dynamics of Norms*. Cambridge University Press.

Bicchieri, C., & Rovelli, C. (1995). Evolution and Revolution: The Dynamics of Corruption. *Rationality and Society*, 7(2), 201–224.

## Week 6 – Evolutionary Game Theory and Cooperation I

### Lecture 11

Evolutionary game theory uses the means of game theory to model evolutionary dynamics. Instead of preference satisfaction evolutionary game theory takes fitness as the pay-off in games. To prepare our engagement with evolutionary game theory we look at the famous tit-for-tat strategy. The expectation of repeated encounters fosters a pro-cooperative tendency, but how exactly?

### Lecture 12

What is it for a strategy to be cooperative in an evolutionary context? Jonathan Birch offers an answer to this question and illustrates it with examples from biological research. We distinguish altruistic cooperation from mere mutualism.

### Mandatory Reading

Axelrod, R., & Hamilton, W. D. (1981). The Evolution of Cooperation. *Science*, 211(4489), 1390–1396.

Birch, J. (2017). *The Philosophy of Social Evolution*. Oxford University Press. Chapter 1. pp. 13-36.

### Suggested Reading

Axelrod, R. (2006). *The Evolution of Cooperation*. Basic Books.

Binmore, K. G. (1994). *Game Theory and the Social Contract. Vol. 1 Playing Fair*. MIT Press.

Binmore, K. G. (1998). *Game Theory and the Social Contract. Vol. 2 Just Playing*. MIT Press.

Gintis, H. (2009). *Game Theory Evolving: A Problem-Centered Introduction to Modeling Strategic Interaction* (2nd ed). Princeton University Press.

Birch, J. (2017). *The Philosophy of Social Evolution*. Oxford University Press.



## **Week 7 – Evolutionary Game Theory and Cooperation II**

### *Lecture 13*

We look closer at the work on indirect reciprocity and read a review paper by Martin Nowak and Karl Sigmund on this topic. They see reputation as a central mechanism for cooperation and as an important step in developing complex social interactions.

### *Lecture 14*

Brian Skyrms has produced some of the most engaging work on evolutionary game theory from a philosophical perspective. We read the first chapter of his book *Evolution of the Social Contract* and learn what including evolution contributes beyond the equilibria solution we are already familiar with.

### **Mandatory Reading**

Nowak, M. & Sigmund, K. (2005). Evolution of indirect reciprocity. *Nature* 437: 1291–98.

Skyrms, B. (1996). *Evolution of the Social Contract*. Cambridge University Press. Chapter 1. pp. 1-21.

### **Suggested Reading**

Boyd, R., & Richerson, P. J. (2005). *The Origin and Evolution of Cultures*. Oxford University Press.

Nowak, M. A. (2006). *Evolutionary Dynamics: Exploring the Equations of Life*. Belknap Press of Harvard University Press.

Nowak, M. A. (2006). Five Rules for the Evolution of Cooperation. *Science*, 314(5805), 1560–1563.

## Week 8 – Evolutionary Game Theory and Cooperation III

### *Lecture 15*

So far, our models neglected the role of location for cooperation. As we have already learned repeated interaction makes a significant difference for whether agents cooperate. Location is an important variable governing repetition. You meet your neighbours more often than a random member of the human population.

### *Lecture 16*

How does evolutionary game theory relate to social structure? Skyrms clearly draws a close connection, but even accepting all his models for evolution we can question whether they help us as much to understand human cooperation in society as he thinks.

### **Mandatory Reading**

Skyrms, B. (2004). *The Stag Hunt and the Evolution of Social Structure*. Cambridge University Press. pp. 15-44.

### **Suggested Reading**

Boyd, R., & Richerson, P. J. (2005). *The Origin and Evolution of Cultures*. Oxford University Press.

Nowak, M. A. (2006). *Evolutionary Dynamics: Exploring the Equations of Life*. Belknap Press of Harvard University Press.

Nowak, M. A. (2006). Five Rules for the Evolution of Cooperation. *Science*, 314(5805), 1560–1563.

Skyrms, B. (1996). *Evolution of the Social Contract*. Cambridge University Press. Chapter 3. pp. 45-62.

## Week 9 – Signalling

### Lecture 17

We return to David Lewis' *Convention* one more time to consider his discussion of signalling. His analysis of common knowledge lays the basis for understanding how actions can take on a conventional meaning. We discuss the limitations this analysis imposes on the concept of signalling.

### Lecture 18

Brian Skyrms has offered a more detailed theory of signalling drawing on evolutionary game theory in addition to Lewis' contribution. We consider this extension to Lewis' proposal and address the differences. No longer do we need common knowledge of the kind introduced by David Lewis.

### Mandatory Reading

Lewis, D. (2011). *Convention: A Philosophical Study*. Blackwell. Chapter 4. pp. 122-159.

Skyrms, B. (2010). *Signals: Evolution, Learning, & Information*. Oxford University Press. Chapter 1 & 2. pp. 5-32.

### Suggested Reading

Skyrms, B. (2010). *Signals: Evolution, Learning, & Information*. Oxford University Press. Chapter 3-5. pp. 33-72.

Godfrey-Smith, P., & Sterelny, K. (2016). Biological Information. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University. Retrieved from <https://plato.stanford.edu/archives/sum2016/entries/information-biological/>

Zahavi, A. (1975). Mate Selection—A Selection for a Handicap. *Journal of Theoretical Biology*, 53(1), 205–214.

## Week 10 – Preference Change and Cooperation I

### Lecture 19

Dietrich and List offer an ambitious approach to preference change. They have developed a general framework for modelling preferences based on motivationally salient properties. Their technical apparatus allows to model a change from purely self-interested to other-regarding preferences. A new model of cooperation becomes possible.

### Lecture 20

We look at a model of preference change, which I propose drawing on the work by Dietrich & List and the theory of signalling we encountered in week 8. Human agents do not change their preferences in an isolated manner, but often communicate while undergoing motivational change. That we can align our preferences makes an especially deep cooperation possible.

### Mandatory Reading

Dietrich, F., & List, C. (2013). Where Do Preferences Come from? *International Journal of Game Theory*, 42(3), 613–637.

Strohmaier, D. (Under Review). Deep Cooperation.

### Suggested Reading:

Dietrich, F., & List, C. (2011). A Model of Non-Informational Preference Change. *Journal of Theoretical Politics*, 23(2), 145–164.

Dietrich, F., & List, C. (2013). A Reason-Based Theory of Rational Choice. *Noûs*, 47(1), 104–134.

Dietrich, F., & List, C. (2016). Reason-Based Choice and Context-Dependence: An Explanatory Framework. *Economics & Philosophy*, 32(2), 175–229.

Grüne-Yanoff, T., & Hansson, S. O. (Eds.). (2009). *Preference Change: Approaches from Philosophy, Economics and Psychology*. Springer.

## Week 11 – Team Reasoning I

### Lecture 21

The approaches to game theory and cooperation we have encountered so far mostly operated on the level of the individual. A number of contributions have questioned this limitation. Robert Sugden (1993) has proposed that individuals reason as teams. We discuss Sugden's argument why altruistic preferences don't suffice to explain our cooperation.

### Lecture 22

Sugden's 1993 paper proposes a largely informal treatment of team reasoning. We look further at how one might formalise this approach. While we are at it we are also considering the different formulation of team reasoning offered by Bacharach.

### Mandatory Reading

Sugden, R. (1993). Thinking as a Team: Towards an Explanation of Nonselfish Behavior. *Social Philosophy and Policy*, 10(1), 69–89.

Gold, N., & Sugden, R. (2007). Theories of Team Agency. In F. Peter & H. B. Schmid (Eds.), *Rationality and Commitment* (pp. 280–312). Oxford University Press.

### Suggested Reading

Bacharach, M. (1999). Interactive Team Reasoning: A Contribution to the Theory of Co-operation. *Research in Economics*, 53(2), 117–147.

Bacharach, M. (2006). *Beyond Individual Choice: Teams and Frames in Game Theory*. Princeton University Press.

Sugden, R. (2000). Team Preferences. *Economics and Philosophy*, 16(02), 175–204.

Sugden, R. (2003). The Logic of Team Reasoning. *Philosophical Explorations*, 6(3), 165–181.

## Week 12 – Team Reasoning II

### Lecture 23

We look beyond the game theoretical debate about team reasoning to include the debate about joint action. Already the paper Gold & Sugden paper from the previous week addressed joint action, but this time we look into it more deeply. I introduce the main arguments made in the debate about what it means for us to act jointly, for example to go on a walk together.

### Lecture 24

Elisabeth Pacherie has offered a more recent discussion about the relation of team reasoning and joint action. She argues that the debate on team reasoning can help to find a compromise between minimalist and maximalist conceptions of joint action. She draws on Bacharach's framing account of team reasoning, which we also consider in this lecture. We end our course by learning what it meant to have engaged in it together!

### Mandatory Reading

Pacherie, E. (2011). Framing Joint Action. *Review of Philosophy and Psychology*, 2(2), 173–192.

### Suggested Reading

Bratman, M. (2014). *Shared Agency: A Planning Theory of Acting Together*. Oxford University Press.

Bratman, M. (1993). Shared Intention. *Ethics*, 104(1), 97–113.

Gilbert, M. (1990). Walking Together: A Paradigmatic Social Phenomenon. *Midwest Studies in Philosophy*, 15(1), 1–14.

Gold, N., & Sugden, R. (2007). Collective Intentions and Team Agency. *The Journal of Philosophy*, 104(3), 109–137.

Searle, J. (1990). Collective Intentions and Actions. In P. R. C. J. Morgan & M. Pollack (Eds.), *Intentions in Communication* (pp. 401–415). MIT Press.

## Essay Topics

Students are encouraged to choose their own essay topics, however, for guidance I propose the following research question, which are available as well:

- Are all social norms cooperative? Discuss potential counterexamples and be clear about the notion of cooperation at play!
- Which kinds of social norms and conventions require common knowledge. Why?
- Signalling lies at the meeting point between evolutionary and psychological game theory. Discuss the differences between different types of signalling with an eye to cooperation.
- We discuss the connection between the team-reasoning and the joint action debate. Is there also a connection between the social norm and the joint action debate? Argue for or against such a connection.
- Does the debate on team-reasoning have anything to contribute to evolutionary game theory? And vice versa? Discuss an example at length.
- Do social norms lead to preference change? Which kinds of preference change? Make sure to clarify you what type of social norms you are writing about!
- Create a plausible formal model of our cooperation in this course and discuss its limitations!

It has been my experience that students who discuss their essay plans with me *at length* receive better grades. Despite this being anecdotal evidence and the causal connection being unclear, I strongly advise you to make use of my office hours!