

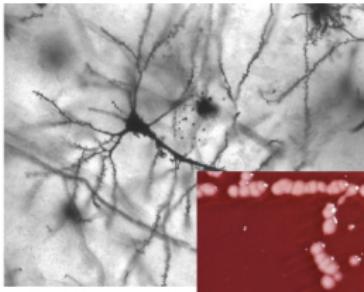
# Game Theory and the Evolution of Signaling

Kevin J.S. Zollman

Carnegie Mellon University

Formal Epistemology Workshop

# Signaling is everywhere



# Philosophical/Scientific questions

- ① Why is signaling stable?
- ② How can reliable signaling come to emerge?
  - Under what conditions?
  - What learning/evolutionary processes?
  - Is there a difference in efficiency among different situations/processes?
- ③ What is the correct methodology?
- ④ What constitutes signaling?
  - How can it be distinguished from other phenomena?
- ⑤ Is there continuity (or similarity) between human language and this type of signaling?
  - Do animal signals have meaning?
  - Can animals have {grammar, pragmatic distinctions, reference, etc.}?
  - Is there deception?

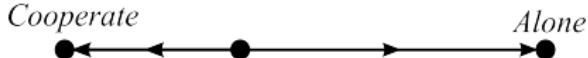
# The static methodology



- Construct a game
- Find its equilibria

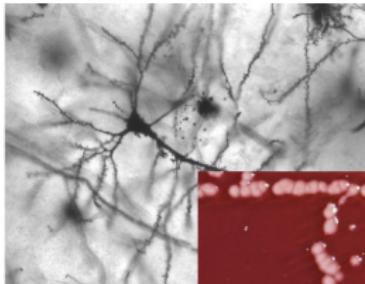
	<i>Cooperate</i>	<i>Alone</i>
<i>Cooperate</i>	3 3	2 0
<i>Alone</i>	0 2	2 2

# The dynamic methodology

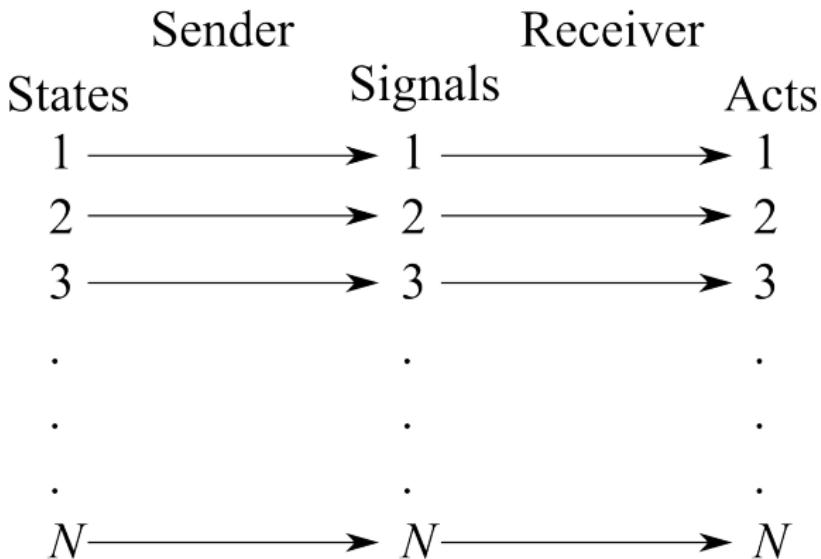


- Construct a game
- **Describe a mechanism for strategy revision**
- Find the long run outcomes of this system

# Signaling is everywhere

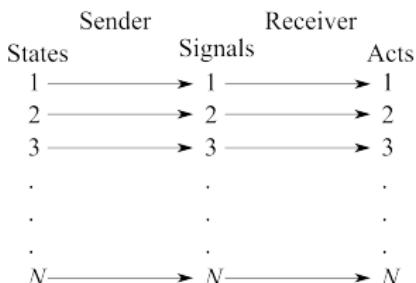


# Lewis signaling game



# Multiple potential end points

- Signaling systems [23]
  - Unique ESS [40]
- Total pooling equilibria [40]
- Partial pooling equilibria [38]
- Which are likely end points of evolution or learning?



# Dynamic approaches

- Population dynamics (evolutionary processes)
  - Replicator dynamics [37]
  - Replicator mutator dynamics [14]
  - Frequency dependent Moran process [26]
  - Spatial evolution [30]
- Learning dynamics (individual learning)
  - Win-stay/Lose-switch or Win-stay/Lose-randomize [31], [27], [5]
  - Imitation dynamics [8]
  - Herrnstein reinforcement dynamics (and variations) [11], [32]
  - Best response dynamics

# A special case

- Two states, two signals, two acts; states equiprobable
- Signaling is guaranteed to emerge in the replicator dynamics [18]
  - Other evolutionary models seems similar [29], [44]
- Signaling is guaranteed in Herrnstein reinforcement learning [1], [35]

# Evolutionary models (general case)

- Without mutation [20], [19], [29]
  - Sometimes perfect languages evolve
  - In some cases no information emerges
  - Sometimes imperfect languages evolve
- With mutation [15], [16], [19], [28]
  - Inefficiencies reduced, but not eliminated
- With population structure [39], [44]
  - Helps, details to follow...

# Learning (general case)

- Win-stay/Lose-switch and Win-stay/Lose-randomize [5]
  - Extremely unstable to errors
- Herrnstein reinforcement learning – like replicator dynamics [1], [2], [35]
- Persistent randomness helps [5]
  - Forgetting the past
  - Persistent shocks
- Learning and evolution interact to improve efficiency [43]
- High rationality dynamics – little is known

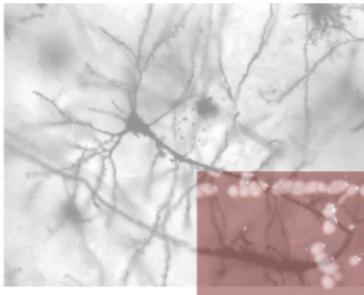
# Extending Lewis' game

- Additional actions or information [19], [45]
  - Indicative/Imperative
- Multiple senders and/or receivers [3], [4], [34], [35]
  - Syntax
  - Natural kinds
  - Information transmission

# Philosophical/Scientific questions

- ① Why is signaling stable?
- ② How can reliable signaling come to emerge?
  - Under what conditions?
  - What learning/evolutionary processes?
  - Is there a difference in efficiency among different situations/processes?
- ③ What is the correct methodology?

# Conflict of interest



# Three canonical cases

- ① Signaling between relatives
- ② Signaling between potential mates
- ③ Signaling between potential rivals

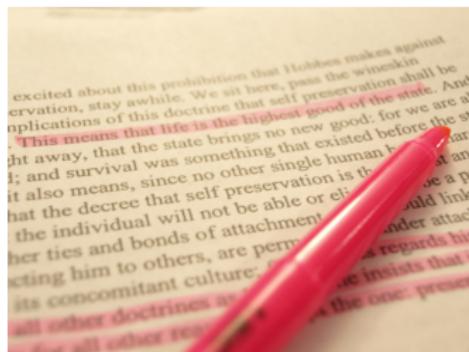
# The problem of honesty...



- All these circumstances feature some conflict of interest
- Here the question of stability has bite
- The standard answer is cost

# Spence and Zahavi

- Why major in philosophy?
- Why does it function as a signal? [36]
  - Cost varies with type
- Similar problems/answers in biology [41], [42]



©C.J. Richey

# Grafen and Hurd

- Cost need not vary with type, benefit can too [13]
- Gain from lying outweighed by cost [17]



© Yoki

# Chick begging

- The phenomena of chick begging
- This can generate a conflict of interest
- Why should communication be honest?



©Idaj

# The game



- Two players: the parent and the chick
- The chick has two states of need *healthy* or *needy*
- The chick has four strategies: beg (or not) dependent on state
- The parent has four strategies: transfer some good to the chick (or not) dependent on the chick begging (or not)

# The game



- Begging may entail some energetic cost
- This is often the explanation for stability in situations of conflict [24], [25], [33]

# A few problems

- Experiments rarely demonstrate significant cost [33]
- Non-signaling equilibria may be Pareto-superior [6], [22], [7]
- It can be very hard to evolve to signaling equilibria
- An odd explanation
  - The stability question is answered...
  - ... at the cost of introducing a mystery to the evolutionary question.
- A few potential alternatives
  - No conflict of interest
  - Partial pooling [6], [22], [7]
  - Spatial structure [21]

# Mate signaling, three stories

- Sensory manipulation [10]
- Sexual selection (the sexy son) [12]
- Signaling

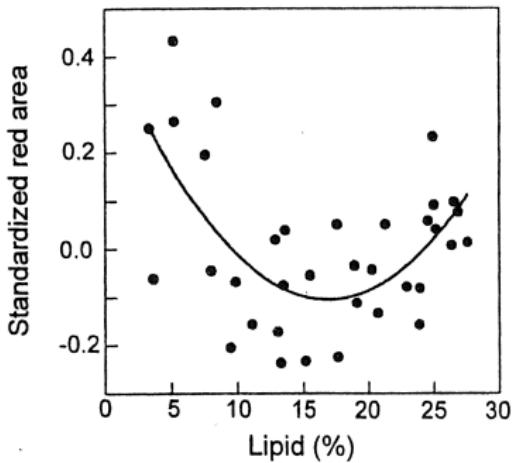
# Coloration in fish



- Could be sensory manipulation
- Could be sexual selection
- Could be signaling

# Deception?

- Low quality males signal as much as high quality males [9]
- Appears to be a case of “bluffing”
- Similar to mimicry



From [9]

# Philosophical/Scientific questions

- ① Why is signaling stable?
- ② How can reliable signaling come to emerge?
  - Under what conditions?
  - What learning/evolutionary processes?
  - Is there a difference in efficiency among different situations/processes?
- ③ What is the correct methodology?

# More questions...

- ① What constitutes signaling?
  - How can it be distinguished from other phenomena?
- ② Is there continuity (or similarity) between human language and this type of signaling?
  - Do animal signals have meaning?
  - Can animals have {grammar, pragmatic distinctions, reference, etc.}?
  - Is there deception?

# Much to do

- Questions of methodology
- Formal questions
- Empirical questions

# Bibliography

- [1] Raffaele Argiento, Robin Pemantle, Brian Skyrms, and Stas Volkov. Learning to signal: Analysis of a micro-level reinforcement model. *Stochastic Processes and their Applications*, 119(2):373–390, February 2009.
- [2] Jeffrey A. Barrett. Numerical simulations of the Lewis signaling game: Learning strategies, pooling equilibria, and the evolution of grammar. Technical Report MBS-06-09, University of California, Irvine: Institute for Mathematical Behavioral Sciences, 2006.
- [3] Jeffrey A. Barrett. The evolution of coding in signaling games. *Theory and Decision*, 2007.
- [4] Jeffrey A. Barrett. Dynamic partitioning and the conventionality of kinds. *Philosophy of Science*, 74:527–546, 2008.
- [5] Jeffrey A. Barrett and Kevin J.S. Zollman. The role of forgetting in the evolution and learning of language. *Forthcoming in Journal of Experimental and Theoretical Artificial Intelligence*, 2009.
- [6] Carl T. Bergstrom and Michael Lachmann. Signalling among relatives. I. Is costly signalling too costly? *Philosophical Transactions of the Royal Society of London B*, 352:609–617, 1997.
- [7] Carl T. Bergstrom and Michael Lachmann. Signalling among relatives. III. Talk is cheap. *Proceedings of the National Academy of Sciences of the USA*, 95:5100–5105, April 1998.
- [8] J. Björnerstedt and Jörgen Weibull. Nash equilibrium and evolution by imitation. In K. Arrow and E. Colombatto, editors, *Rationality in Economics*. Mcmillan, 1999.
- [9] U. Candolin. The relationship between signal quality and physical condition: Is sexual signaling honest in three-spined stickleback? *Animal Behavior*, 58:1261–1267, 1999.
- [10] Richard Dawkins and J.R. Krebs. Animal signals: information or manipulation. In J.R. Krebs and N.B. Davies, editors, *Behavioral ecology: an evolutionary approach*, pages 282–309. 1978.
- [11] Ido Erev and Alvin E. Roth. Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. *The American Economic Review*, 88(4):848–881, September 1998.
- [12] Ronald A. Fisher. *The Genetical Theory of Natural Selection*. Oxford University Press, Oxford, 1930.
- [13] Alan Grafen. Biological signals as handicaps. *Journal of Theoretical Biology*, 144:517–546, 1990.
- [14] Josef Hofbauer. The selection mutation equation. *Journal of Mathematical Biology*, 23:41–53, 1985.
- [15] Josef Hofbauer and Simon Huttegger. Selection-mutation dynamics of signaling games with two signals. In Anton Benz, Christian Ebert, and Robert von Roij, editors, *Proceedings of the ESSLLI 2007 Workshop on Language, Games, and Evolution*, August 2007.
- [16] Josef Hofbauer and Simon Huttegger. Feasibility of communication in binary signaling games. *Journal of Theoretical Biology*, 254(4):843–849, 2008.
- [17] Peter L. Hurd. Communication in discrete action-response games. *Journal of Theoretical Biology*, 174:217–222, 1995.
- [18] Simon Huttegger. Evolution and explanation of meaning. *Philosophy of Science*, 74(1):1–27, January 2007.
- [19] Simon Huttegger. Evolutionary explanations of indicatives and imperatives. *Erkenntnis*, 66:409–436, 2007.
- [20] Simon Huttegger, Brian Skyrms, Rory Smead, and Kevin Zollman. Evolutionary dynamics of Lewis signaling games: Signaling systems, partial pooling. *Forthcoming in Synthese*, 2009.
- [21] David C. Krakauer and Mark Pagel. Spatial structure and the evolution of honest cost-free signaling. *Proceedings of the Royal Society of London B*, 266:365–373, 1995.
- [22] Michael Lachmann and Carl T. Bergstrom. Signalling among relatives. II. Beyond the tower of babel. *Theoretical Population Biology*, 54:146–160, 1998.
- [23] David Lewis. *Convention: A Philosophical Study*. Harvard University Press, Cambridge, 1969.
- [24] John Maynard Smith. Honest signalling, the Philip Sidney game. *Animal Behavior*, 42:1034–1035, 1991.
- [25] John Maynard Smith and David Harper. *Animal signals*. Oxford University Press, Oxford, 2003.
- [26] P.A.P. Moran. *The Statistical Process of Evolutionary Theory*. Clarendon Press, Oxford, 1962.
- [27] Martin Nowak and Karl Sigmund. A strategy of win-stay, lose-shift that outperforms tit-for-tat in the Prisoner's Dilemma game. *Nature*, 364:56–58, July 1 1993.
- [28] Christina Pawlowsits. Finite populations choose an efficient language. *Journal of Theoretical Biology*, 249:606–617, 2007.
- [29] Christina Pawlowsits. Why evolution does not always lead to an optimal signaling system. *Games and Economic Behavior*, 63:203–226, 2008.
- [30] Gregory B. Pollock. Evolutionary stability of reciprocity in a viscous lattice. *Social Networks*, 11:175–212, 1989.
- [31] Herbert Robbins. Some aspects of the sequential design of experiments. *Bulletin of the American Mathematical Society*, 58:527–535, 1952.
- [32] Alvin E. Roth and Ido Erev. Learning in extensive-form games: Experimental data and simple dynamics models in the intermediate term. *Games and Economic Behavior*, 8:164–212, 1995.
- [33] William A. Searcy and Stephen Nowicki. *The Evolution of Animal Communication*. Princeton University Press, Princeton, 2005.
- [34] Brian Skyrms. Evolution of signaling systems with multiple senders and receivers. *Philosophical Transactions of the Royal Society of London B*, 364:771–779, 2009.
- [35] Brian Skyrms. *Signals: Evolution, Learning, and the Flow of Information* (book manuscript). 2009.
- [36] Michael Spence. Job market signaling. *The Quarterly Journal of Economics*, 87(3):355–374, 1973.
- [37] P. Taylor and L. Jonker. Evolutionarily stable strategies and game dynamics. *Mathematical Biosciences*, 40:145–156, 1978.
- [38] Peter E. Trapa and Martin A. Nowak. Nash equilibria for an evolutionary language game. *Journal of Mathematical Biology*, 41:172–188, 2000.
- [39] Elliot Wagner. Communication and structured correlation. *Forthcoming in Erkenntnis*, 2009.
- [40] Karl Wärneryd. Evolutionary stability in unanimity games with cheap talk. *Economics Letters*, 36:375–378, 1991.
- [41] Amotz Zahavi. Mate selection – a selection for a handicap. *Journal of Theoretical Biology*, 53:205–214, 1975.
- [42] Amotz Zahavi and Avishag Zahavi. *The Handicap Principle: A Missing Piece of Darwin's Puzzle*. Oxford University Press, New York, 1997.
- [43] Kevin Zollman and Rory Smead. Plasticity and language: an example of the Baldwin effect? *Philosophical Studies*, forthcoming, 2009.
- [44] Kevin J.S. Zollman. Talking to neighbors: The evolution of regional meaning. *Philosophy of Science*, 72:69–85, 2005.
- [45] Kevin J.S. Zollman. Separating directives and assertions using simple signaling games. *Manuscript*, 2009.