On Agent-Based Modeling of Complex Systems: Learning and Bounded Rationality

Sanmay Das

Department of Computer Science and Engineering
University of California at San Diego
La Jolla, CA, 92093-0404

Abstract

This paper seeks to connect the literatures from artificial intelligence, economics, and cognitive science to make the case that not only is the notion of bounded optimality from the AI literature the right goal for agent design, it can also serve as a principled means for modeling boundedly rational agents in complex systems like economic markets. While appealing, this goal leaves open two critical questions. First, bounded optimality is defined over an expected set of problems the agent might face and it is not obvious what criterion to use for expected problems. Second, it will typically be impossible to design a provably boundedly optimal agent even given a set of expected problems, because the agent design problem itself is intractable. These problems become particularly important when agents must learn from their environments. In order to deal with these questions, we may need to abandon the formalism of mathematics and instead look towards the process of science and engineering. I argue that it is critical to evaluate agents in terms of the expected set of problems they would face if they were deployed in the real world, either in software or in hardware, and the agent programs we use for modeling should be the best known program for any given problem subject to the broader expected set of problems the algorithm might be expected to solve – the algorithm we would choose to use if we had to hand over control of our own behavior in that domain to an artificial agent.

Key words:
PACS:

Email address: sanmay@cs.ucsd.edu (Sanmay Das).

1 Introduction

Human beings regularly make extremely complex decisions in a world filled with uncertainty. For example, we decide which gas station to refuel at, and whether to put money into a bank account or a retirement fund, with surprisingly little effort. Designing algorithms for artificial agents in similar situations, however, has proven extremely difficult. Historically, research in artificial intelligence has focused on designing algorithms with the ability to solve these problems in principle, given infinite computational power [Russell, 1997]. Many problems that arise in everyday decision-making are likely to be impossible to solve perfectly given computational constraints, so this kind of *calculative rationality*, as Russell calls it, is not a particularly interesting practical goal. In order to progress towards designing an intelligent system, we need a better theory of decision-making and learning when rationality is bounded by computational resource considerations. Not only is such a theory critical to our understanding of the nature of intelligence, it could also be applied to study interaction between agents in increasingly realistic models of social and economic systems by serving as a replacement for the classical economic theory of unbounded rationality. Russell [1997] discusses two possible theories, *metalevel rationality* and *bounded optimality*. Russell and Norvig [2003] also consider a more general notion of bounded rationality in the tradition of Herbert Simon's (1955) satisficing. Bounded optimality is in many ways the most appealing of these theories, since it strives to replace agents that always make rational *decisions* with rational *agent programs*. Bounded optimality could also be a more plausible model of how humans work. We know that human beings do not make rational decisions, but it is possible that our brains have evolved to become an optimal or near-optimal decision-making system for the environments we inhabit.

One of the great attractions of bounded optimality is that it can serve both as a definition of intelligence meeting the needs of research in AI and as a formal model of decision-making to replace unbounded rationality. Unfortunately, while bounded optimality might be the right goal to strive for in agent design, achieving this goal, or even knowing whether it can be achieved, is difficult when agents are uncertain about the world. This issue becomes particularly important in the context of agents that are not fully informed about the environment they are placed in, and who must learn how to act in a successful manner. The question of what is meant by boundedly rational learning and how it can be analyzed is a difficult one that is discussed in some detail below.

The broad plan of this paper is as follows.

- I start by considering various perspectives on bounded rationality. While Russell [1997] and Parkes [1999] provide plenty of detail on most of the perspectives related to agent design, I will consider in more detail the tradition of bounded rationality research started by Herbert Simon with the notion of satisficing, and continued in the program of fast and frugal heuristics by Gerd Gigerenzer and others [Gigerenzer and Goldstein, 1996, Gigerenzer and Selten, 2001, *inter alia*].

- After that, I move on to considering how the field of machine learning, and in particular, reinforcement learning, needs to adapt in order to move towards the goal of designing boundedly optimal agents for increasingly complex domains.
- In the last major section, I discuss how a theory of boundedly optimal agents may provide a compelling replacement for the unboundedly rational agents of economic theory. The notion of bounded optimality can serve as a principled approach to modeling complex systems and address many of the criticisms of bounded rationality research found in the economics literature.

2 Bounded Rationality and Agent Design

The agent-based approach to AI involves designing agents that "do the right thing" [Russell, 1997]. The "right thing" in this context means that agents should take actions that maximize their expected utility (or probability of achieving their goals). Ideally, an agent should be perfectly rational. Unfortunately, there is really no such thing as a perfectly rational agent in the world. As Russell [1997, pp. 6] says, "physical mechanisms take time to process information and select actions, hence the behavior of real agents cannot immediately reflect changes in the environment and will generally be suboptimal." *Calculative rationality*, the ability to compute the perfectly rational action in principle, given sufficient time and computational resources, is not a useful notion, because agents that act in the world have physical constraints on when they need to choose their actions. We are left to contemplate other options for agent design.

The heuristics and biases program made famous by Kahneman and Tversky studies actual human behavior and how it deviates from the norms of rational choice. This program is not in any way prescriptive, as it mainly focuses on cataloging deviations from the presumed normative laws of classical decision theory. Thus, this program does not provide any suitable *definitions* for intelligence that we can work towards, although understanding human deviations from decision-theoretic norms might prove informative in the design of good algorithms, as is argued below.

Another approach from the literature of cognitive science is the use of "satisficing" heuristics in the tradition of Simon [1955], who introduced the notion that human decision-makers do not exhaustively search over the space of outcomes to choose the best decision, but instead stop as soon as they see an outcome that is above some satisfactory threshold "aspiration level." Conlisk [1996] cites various papers in the economics literature that start from Simon's notion of bounded rationality, and claims that, within economics, "the spirit of the idea is pervasive." In cognitive science and psychology, Gigerenzer and others have recently popularized the use of "fast and frugal" heuristics and algorithms as the natural successor to satisficing. Gigerenzer and Goldstein [1996] state their view of heuristics as being "ecologically rational" (capable of exploiting structures of information present in the environment) while nevertheless violating classical norms of rationality. They have a program to design computational models of such heuristics, which are "fast, frugal and simple enough to operate effectively when time, knowledge, and computational might are limited" while making it quite clear that they do not agree with the view of heuristics as "imperfect versions of optimal statistical procedures too complicated for ordinary minds to carry out" [Goldstein and Gigerenzer, 2002]. They reject the Kahneman-Tversky program because it maintains the normative nature of classical decision theory.

It is interesting that in their 1997 paper, Gigerenzer and Goldstein held a simulated contest between a satisficing algorithm and "rational" inference procedures, and found that the satisficing procedure matched or outperformed more sophisticated statistical algorithms. The fact that a simple algorithm performs very well on a possibly complex task is not surprising in itself, but what is very clear is that if it can be expected to perform *better* than a sophisticated statistical algorithm, it must either be a better inference procedure or have superior prior information encoded within it *in the context of the environment in which it is tested*. As agent designers, if we had to design an agent that solves a given range of problems, and had access to the information that a satisficing heuristic was the best known algorithm for that range of problems, it would be silly not to use that algorithm in the agent, all else being equal. While I will return to this issue below, the problem with fast and frugal heuristics as a program for agent design is the loose definition of what constitutes a satisfactory outcome, or of what kinds of decision-making methods are "ecologically rational" in the language of Goldstein and Gigerenzer. How do we know that one heuristic is better than another? What if our agent has to perform many different tasks?

Russell proposes two other options for a goal for agent design – *metalevel rationality* and *bounded optimality*. Metalevel rationality involves reasoning about the costs of reasoning. An agent that is rational at the metalevel "selects computations according to their expected utility" [Russell, 1997]. This is what Conlisk [1996] refers to as *deliberation cost*, and both Russell and Conlisk make the explicit connection to the analogous *value of information*. Conlisk argues strongly for incorporating deliberation cost into optimization problems that arise in economics, saying that human computation is a scarce resource, and economics is by definition the study of the allocation of scarce resources. He suggests that instead of optimally solving an optimization problem P, a decision-maker should solve an augmented problem $F(P)$ in which the cost of deliberation is taken into account. The problem, as he realizes, is that it is also costly to reason about $F(P)$, and, therefore, theoretically at least, one should reason about $F(F(P)), F(F(F(P))), \ldots$. This infinite regress is almost always ignored in the literature that does take deliberation cost into account, assuming that $F(P)$ is, in some sense, a "good enough" approximation.

Economics is not alone in this myopic consideration of deliberation cost. In the AI literature, the tradition of studying deliberation cost (or the essentially equivalent concept which Russell calls the *value of computation*) as part of solving a decision problem dates back to at least the work of Eric Horvitz [1987], and almost all the algorithms that have been developed have used myopically optimal metareasoning at the first level, or shown bounds in very particular instances. The history of metalevel "rationality" in the AI literature is more that of a useful tool for solving certain kinds of problems (especially in the development of anytime algorithms) than as a formal specification for intelligent agents. This is mostly because of the infinite regress problem described above – as [Russell, 1997, page 10] writes (of the first metalevel, or the problem $F(P)$ in Conlisk's notation), "perfect rationality at the metalevel is unattainable and calculative rationality at the metalevel is useless."

This leaves us with Russell's last, and most appealing, candidate – *bounded optimality*, first defined by Horvitz [1987] as "the optimization of [utility] given a set of assumptions about expected problems and constraints on resources." Russell [1997] says that bounded optimality involves stepping "outside the agent" and specifying that the agent *program* be rational, rather than every single agent decision. An agent's decision procedure is bounded optimal if the expected utility of an action selected by the procedure is at least as high as that of the action selected by any decision procedure subject to the same resource bounds in the same environment. Of course, we are now requiring a lot of the agent designer, but that seems to make more sense as a one-time optimization problem than requiring the same of the agent for every optimization problem it faces. The problem with bounded optimality is that, for any reasonably complex problem, it will prove incredibly hard to design a boundedly optimal agent, and it might also be hard to prove the optimality of the agent program. We have shifted the burden of rationality to the designer, but the burden still exists.

Nevertheless, bounded optimality seems to be the right goal. If one agent algorithm can be shown to perform better than another given the same beliefs about the set of problems the agent may face and the same computational resources, the first algorithm should be the one used. Of course, this can be problematic because it is not clear *where* it must perform better. What if the prior beliefs of the agents are completely wrong? Again, the only solution is to analyze this from the perspective of the agent designer. Rodney Brooks [1991] argues that the real world is its own best model. While I do not endorse the notion that the task of modeling by simplifying is therefore useless, I do think we should hold boundedly optimal algorithms to the standard of reality. The performance of algorithms should be tested on problems that are as realistic as possible.

Let me use this for a quick foray into the broader question of what we should be doing *as artificial intelligence researchers* when we design algorithms intended to be boundedly optimal. The two points that [Brooks, 1991, page 140] makes about the direction of AI research are (quoting directly):

- We must incrementally build up the capabilities of intelligent systems, having complete systems at each step of the way and thus automatically ensure that the pieces and their interfaces are valid.
- At each step we should build complete intelligent systems that we let loose in the real world with real sensing and real action. Anything less provides a candidate with which we can delude ourselves.

Brooks carries this forward to argue for robotics and immersion in the real *physical* world as the most important research program in AI. Oren Etzioni [1993] argues against this view, arguing that the creation of so-called softbots, agents embedded in software, is perhaps a more useful focus for AI research, since it allows us as researchers to tackle high-level issues more immediately without getting sidetracked. The concept of designing boundedly optimal agents is not necessarily tied to either of these research directions. While the later focus of this paper on modeling complex systems falls more in line with Etzioni's arguments than Brooks', the basic goal of agent design is the same in both areas, and both are useful programs of research in AI. This goal is to solve the engineering problem of creating an agent that will be optimally successful in the environments in which we, as agent designers, expect it to be placed, given the resources available to the agent. The agent could be a software agent attempting to trade optimally in an online market, or it could be a robot attempting to navigate many different kinds of terrain, or almost anything else one can think of.

One critical caveat needs to be mentioned – eventually, we would like to create agents that are generally intelligent across a range of domains and problems (holding a conversation, ordering food from a restaurant, buying a plane ticket, and so on). These agents will have limited computational resources, and therefore, they might not be able to spend a lot of these resources on any one task. They will need to be boundedly optimal in the context of all the problems they have to solve, and therefore, they might have to be designed so that they would not be boundedly optimal for any one of those problems given the same computational resources. This caveat will also apply to modeling economic and social systems. We cannot assume too much in the way of computational resources or exclusivity of access to these resources when we model using the methodology proposed here. In fact, suppose the human brain itself is boundedly optimal for the environment it inhabits with respect to some kind of evolutionary survival utility function (or fitness function). Then the kinds of deviations we see from decision-theoretic norms could be explained by the fact that even if an algorithm were not optimal for a particular class of problems, if the agent could perhaps face a significantly larger class of problems, or a better inference procedure were significantly more costly, it might be boundedly optimal to program the agent with that algorithm since this would save valuable resources that could be devoted to other problems. In fact, some of the algorithms described as "fast and frugal heuristics" [Gigerenzer and Goldstein, 1996, Goldstein and Gigerenzer, 2002] may well be bounded optimal in the larger context of the human brain. Some evidence of how understanding the heuristics we use can be important in the context of algorithm development can be seen in the FoRR architecture which leads to, for example, an effective navigation algorithm [Epstein, 1998], and in the development of a model for how humans effectively solve the traveling salesperson problem Best [2004].

3 Bounded Optimality and Learning

In the tradition of economics and decision theory, rational learning is understood to mean Bayesian updating of beliefs based on observations. This definition finesses two major problems we encounter in agent design.

- It shifts responsibility onto the designer's beliefs about the set of problems the agent might face. It could be very hard to design an appropriate prior for an agent that must make many different kinds of decisions in the world.
- Even if the designer's prior beliefs are correct, full Bayesian learning is computationally infeasible in all but the simplest cases. There may be no single algorithm that provably always outperforms others in selecting optimal actions within the specified computational limits, or finding such an algorithm may be hard.

First, let us consider a simple illustrative problem that is an extension of the classical supervised learning framework to a situation where a predictive agent receives utility from making correct predictions while her actions have no influence on the environment. Later we will turn to the even more difficult case in which the agent's actions impact its environment.

3.1 Learning to Predict:

Consider a situation in which an agent, call her Mary, receives as input a real-valued vector X at each time step, and has to predict Y, where Y is known to be a (probabilistic or noisy) function of X. The two standard cases are regression, where Y is real valued, or classification, where Y takes on one of a few specific different values. For simplicity, consider the binary classification case where $Y \in \{0, 1\}$. Suppose Mary's task is to predict whether Y will be 0 or 1. Immediately after she makes her prediction, she is informed of the true value of Y and receives utility 1 for making the correct prediction, and 0 for making the wrong prediction. Suppose Mary knows that she will see 100 such examples, and her goal is to maximize the utility she receives over the course of this game. How should she play? The theory of unbounded rationality and "rational learning" as used by economics would say that Mary starts with a prior $\Pr(Y|X)$, makes her decision based on the particular instantiation $X = x$ that she sees (predicting $Y = 1$ if $\Pr(Y = 1|X = x) > \Pr(Y = 0|X = x)$ and $Y = 0$ otherwise), and then updates her estimate $\Pr(Y|X)$ using Bayes' rule after the true value of Y is revealed to her. Unfortunately, even ignoring the issue of how to specify a good prior, performing the full Bayesian updates at each step is computationally prohibitive. This brings us into the sphere of bounded optimality. What method would achieve as high a utility as possible for Mary given reasonable computational resources? This question does not have a definite answer, even if we clearly specify the exact resource constraints on the agent. Many different algorithms for the supervised learning problem, both online and offline, could be applied in this situation. Mary could memorize all the examples she sees and use a support vector machine to learn a classifier after each step (even then, she would need to make choices about kernels and parameters), or she could use an ensemble classifier like boosted decision trees, or an online algorithm like the Widrow-Hoff procedure, but there is no one algorithm that is clearly better than the others in terms of expected utility across domains.

3.2 Learning How to Act

An agent that acts in the world gains utility from the actions it takes and also learns about the world through the effects of its actions. It must balance exploration (taking myopically suboptimal actions in order to learn more) with exploitation of what it has learned. Perfectly rational Bayesian learning is only possible for certain families of prior beliefs and extremely simple problems. The discovery of an optimal algorithm that balances exploration and exploitation for even the "simple" case of the multi-armed bandit problem was hailed as almost miraculous when it was first published.

The multi-armed bandit problem [Berry and Fristedt, 1985, Gittins and Jones, 1974, *inter alia*] is often considered the paradigmatic exploration-exploitation problem, because the tradeoff can be expressed simply. A single agent must choose which arm of a slot machine to pull at each time step, knowing that the arms may have different reward distributions. The goal of the agent is to maximize the discounted sum of payoffs. It turns out there is actually an optimal way to play the multi-armed bandit under the assumptions of stationary reward distributions and geometric discounting [Gittins and Jones, 1974]. The optimal action at any time is to play the arm with the highest Gittins index, a single number associated with each arm that is based solely on the history of rewards associated with that arm. The Gittins index is reasonably easy to compute for certain families of reward distributions, but can be difficult in other circumstances.

The multi-armed bandit is an easy problem, though, compared to the kinds of problems faced by agents in realistic environments. For example, it is hard to define "optimality" in a nonstochastic bandit problem, where the reward for each arm at each time period may have been picked by an adversary [Auer et al., 2002], or a nonstationary case, where the reward distribution for each arm may change over time. Or consider the case where there are multiple people playing the bandit and the arms of the bandit themselves have agency and must try to maximize their own reward, which is dependent on who pulls them in each period (the "two-sided" bandit problem [Das and Kamenica, 2005]). Even solving problems with a Markovian structure can be extremely hard, especially when they are not fully observable.

3.3 Evaluating Learning Algorithms

While a whole line of literature stresses the importance of resource bounds in understanding computational limitations on rationality in many decision-making environments, the problem becomes particularly severe when the agent does not have a perfect model of its environment and must learn this model through experience. The scenarios examined above raise two fundamental questions for the field of machine learning which parallel those in the introductory discussion. First, if we are seeking to design a boundedly optimal learning algorithm, what should constitute the expected set of problems against which it is evaluated? Second, even given such a set of problems, what should we do if we cannot find a boundedly optimal algorithm, or prove its optimality? I will defer discussion of the second question to the next section, because it will be important for understanding how we can use the notion of bounded optimality in modeling complex systems.

Of course it is impossible to definitively answer these questions, but it is important to keep them in mind as researchers. It is critical to evaluate algorithms from the perspective of performance in the real world, given the expected set of problems an agent would face if it were deployed in the world, either physically or as a software agent. While there is of course value to the traditional computer science program of proving worst case bounds and evaluating algorithms on arbitrary problem spaces, at some stage (not necessarily at the very initial stage of development) algorithms that are designed to be parts of intelligent agents must face the discipline of the real world.

The problem of learning how to act when an agent gets rewards from its interactions with the environment has been studied extensively (especially in the Markovian framework) in the reinforcement learning and neurodynamic programming literatures. However, with some exceptions (see, for example, [Kearns and Singh, 2002]), analysis has typically focused on the ability of algorithms to eventually learn the true underlying model, and hence the asymptotically optimal decision procedure, and not on the expected utilities achieved by algorithms in potentially limited interactions with the environment. This expected utility viewpoint becomes especially important in the true agent design problem, because our goal must be to design agents that can take the actual costs of exploration and learning into account. An agent does not have access to an offline model of its environment so that it can improve its performance before acting in the world. The agent must be able to make tradeoffs in an online manner, where failure or poor performance immediately impacts the agent negatively. It is therefore critical from the perspective of AI research to attack problems of learning how to act from the perspective of expected utility received in the world, especially keeping in mind the real costs of exploration. It will probably prove natural to adopt a Bayesian perspective in analyzing this issue. To quote Berry and Fristedt [1985], in their discussion of bandit problems, "[it] is not that researchers in bandit problems tend to be 'Bayesians'; rather, Bayes's theorem provides a convenient mathematical formalism that allows for adaptive learning, and so is an ideal tool in sequential decision problems."

4 Bounded Optimality and Modeling Complex Systems

Historically, research on the outcomes of interaction between self-interested optimizing agents has been the domain of economic theory. Economists place a high value on analytical tractability and model parsimony. They tend to simplify models until agent behavior and interactions can be reduced to a set of equations that provide intuition to the person analyzing the system. In the words of Brian Arthur [1999], "conventional economic theory chooses not to study the unfolding of the patterns its agents create, but rather to simplify its questions in order to seek analytical solutions." These simplified questions have no need of the notion of bounded optimality, because the decision (and learning) problems faced by agents are "easy" in the sense that they can be solved efficiently without using excessive computational resources. If we move to more complicated models, we will necessarily have to examine more difficult agent decision problems, and we need to think about agent decision-making differently. In a seminal paper, Herbert Simon [1955] identified the problem:

> Broadly stated, the task is to replace the global rationality of economic man with a kind of rational behavior that is compatible with the access to information and the computational capacities that are actually possessed by organisms, including man, in the kinds of environments in which such organisms exist. One is tempted to turn to the literature of psychology for the answer. Psychologists have certainly been concerned with rational behavior, particularly in their interest in learning phenomena. But the distance is so great between our present psychological knowledge of the learning and choice processes and the kinds of knowledge needed for economic and administrative theory that a marking stone placed halfway between might help travelers from both directions to keep to their courses.

How about bounded optimality as a milestone? With the increasing availability of huge amounts of computational power on every researcher's desktop, we can now analyze models using computational tools, and so it is no longer absolutely critical to simplify models to the extreme. This means we can study models in which agent decision problems are no longer easy and computational resource constraints must be taken into account. The literature in the broad field of "complexity science," and particularly econophysics, with its focus on financial markets, has already taken the lead in this direction, and Farmer et al. [2005] provide an overview of research from that perspective, while raising many of the same questions about modeling methodology that this paper seeks to answer. But while econophysics typically seeks to build up from the opposite extreme, random or near-random agent behavior, while the dynamics of interaction remain complex, the approach from AI should consider more complex models of agent decision-making. The ability to study more complex models is particularly important, as noted above, in circumstances where agents are not perfectly informed of the structure or the parameters of the environments in which they are placed. How should this kind of modeling proceed so as not to fall into the traps that have made bounded rationality research anathema to many mainstream economists?

John Conlisk [1996] both raises and answers many of the criticisms of bounded rationality research. Perhaps the most important and frequent such criticism is that the decision procedures modeled by those conducting the research are ad hoc. Conlisk summarizes this argument as follows "Without the discipline of optimizing models, economic theory would degenerate into a hodge podge of ad hoc hypotheses which cover every fact but which lack overall cohesion and scientific refutability" [pp. 685]. He goes on to say that discipline comes from good scientific practice, not strict adherence to a particular approach, and suggests that modeling bounded rationality based on deliberation cost would enforce a certain discipline. Conlisk's preferred approach is the equivalent of metalevel rationality, but I would venture to propose that bounded optimality might be both a better model for the purpose of enforcing discipline (given the infinite regress problem) and a more satisfying model of the processes of actual decision-makers in the world. Maybe humans and economic firms *do* take the best actions available, given their capacities for reasoning about these actions and knowledge of the environment.

However, from the discussion above, it is clear that very often we will not know if an agent algorithm is in fact boundedly optimal or not. Does this invalidate the idea that bounded optimality can serve as a principled replacement for unbounded rationality in economic models? I would argue that this is not the case, but we have to turn to good practice in science and engineering rather than relying on the formalism of mathematics. We should strive to engineer good algorithms for complicated problems, attempting to be rational ourselves in the design of these algorithms. The choice of how to model an agent that is part of a complex system should be made by using the same agent that we would use if we had to engineer an agent for maximally successful performance in that system given our beliefs about the system. Hopefully we could eventually reach more and more realistic models, benefiting both algorithm development and hence the eventual goal of building an intelligent agent as well as our understanding of complex social and economic systems. Let us examine a case study.

4.1 Learning in Economic Models

Many economic models already account for learning in an explicit manner. There is nothing particularly novel about suggesting that agents learn from the environment around them. In an interview with Thomas Sargent, Evans and Honkapohja [2005] explore many of the issues related to how the program called "learning theory" originated in the macroeconomics literature. Let me briefly summarize Sargent's overview of the development of this literature in order to draw a parallel to the overall argument of this paper.

Learning theory in macroeconomics originated in some ways as a response to rational expectations economics. In the rational expectations literature, there exists what Sargent calls a "communism of models" in which all the agents share the same model of the world, and this is the true model, or "God's model".[1] There is no place for different beliefs in rational expectations the-

[1] Sargent uses the term "model" to mean probability distributions over all the

ory. Margaret Bray and David Kreps started a research program to show what would happen if you endowed agents with different beliefs, learning algorithms, and data on what had happened in the past. Agents should continue to update their models and then optimize based on their current beliefs. The interesting outcome was that in many cases, the only possible outcomes of the system were close to rational expectations equilibria. In some cases, the learning models helped to eliminate possible rational expectations equilibria because they could not be reached through the learning dynamic. Further research along these lines has started to examine what happens to equilibria when agents can have model uncertainty, not just parameter uncertainty. Along with equilibrium selection, the theory has also contributed much by helping to understand the rates of reaching equilibrium in different problems and in characterizing the situations where the system dynamics show serious deviations from expected equilibrium behavior. Similar research has now become an integral part of the game theory literature as well [Fudenberg and Levine, 1998].

How does this relate to bounded rationality? Learning behavior is not necessarily "irrational," so why would we have to think of it any differently? Sargent echoes this manner of thinking in talking about "robust control" (in which an agent explicitly has doubts about her model of the world and must take these into account in decision-making) when he says that it is not a type of bounded rationality because "[the agent's] fear of model misspecification is out in the open" which makes her smarter than a rational expectations agent [Evans and Honkapohja, 2005]. The problem that arises with this belief is that it is *impossible* to do provably optimal learning in a model in which there is any kind of complexity to the agents' beliefs. We must take computational resource constraints into account, and agents cannot be unboundedly rational. It becomes very hard to even define rationality meaningfully in most of these situations. To quote Horvitz [1987], "[Constraints in resources] can transform a non-normative technique into the 'preferred choice' of devout Bayesians, and can convert the strictest formalists into admirers of heuristics." We have to think about what constitutes a "good" learning algorithm and whether it makes sense (both from the agent-design and the modeling perspectives) to endow the agent with that algorithm. The learning literature in economics typically focuses on very simple methods of least-squares learning which might not be the choice we would make as agent designers if we had to write an algorithm to participate in the world we are modeling.

inputs and outputs of the larger economic model.

5 Conclusion

This paper attempts to link the literatures of artificial intelligence, economics, and cognitive science so that those familiar with any of these disciplines will be able to see the parallels and connections easily. I have argued for a particular methodology for both designing artificial agents and for modeling agents that are participants in complex systems like economic markets. This methodology is to start by assuming bounded optimality, or the rationality of the agent program, as the goal in both cases. Since this goal needs to be defined in terms of the expected set of problems an agent will face, we should design agents that would perform successfully in the real world, and expect that the set of problems the agent will face is the set of problems it would encounter in the world. Finally, we will not necessarily ever know or be able to show that an agent we have designed is boundedly optimal. We might have to replace our desire for proving that an agent *is* boundedly optimal with a more scientific or engineering based approach, in which we try to design the best algorithm *so far developed* for a problem given the computational and other constraints on the algorithm.

To conclude with an example, suppose we were to design an agent that took care of our finances. We would want it to successfully trade stocks and bonds, perhaps even foreign exchange, while at the same time taking care of more mundane tasks like maintaining bank accounts and paying mortgages. It is not far-fetched to think that we will in the near future be able to design an agent that is close to boundedly optimal for this problem. Eventually we might be able to use insights from the design of this financial agent to build a truly intelligent agent, but in the meanwhile, if we are happy deploying the agent to take care of our finances, we should use it as our model of an economic decision-making agent when we model financial markets.

Acknowledgements

This paper is based largely on work done while I was a Ph.D student at the MIT Center for Biological and Computational Learning. I am grateful to Ranen Das, Jennifer Dlugosz, Leslie Kaelbling, Emir Kamenica, Adlar Kim, Andrew Lo, Sayan Mukherjee, Tomaso Poggio and John Tsitsiklis for discussions that shaped my thinking on these issues. This research was supported by an MIT Presidential Fellowship and by grants from Merrill-Lynch, National Science Foundation (ITR/IM) Contract No. IIS-0085836 and National Science Foundation (ITR/SYS) Contract No. IIS-0112991. Additional support was provided by the Center for e-Business (MIT), DaimlerChrysler AG, Eastman Kodak Company, Honda R&D Co., Ltd., The Eugene McDermott

Foundation, and The Whitaker Foundation.

References

W. Brian Arthur. Complexity and the economy. *Science*, 284:107–109, 1999.

Peter Auer, Nicolò Cesa-Bianchi, Yoav Freund, and Robert E. Schapire. The nonstochastic multiarmed bandit problem. *SIAM Journal on Computing*, 32(1):48–77, 2002.

D.A. Berry and B. Fristedt. *Bandit Problems: Sequential Allocation of Experiments*. Monographs on Statistics and Applied Probability. Chapman and Hall, London, UK, 1985.

Bradley J. Best. *Modeling Human Performance on the Traveling Salesperson Problem: Empirical Studies and Computational Simulations*. PhD thesis, Carnegie Mellon University, 2004.

Rodney A. Brooks. Intelligence without representation. *Artificial Intelligence*, 47:139–159, 1991.

John Conlisk. Why bounded rationality? *Journal of Economic Literature*, 34 (2):669–700, 1996.

Sanmay Das and Emir Kamenica. Two-sided bandits and the dating market. In *Proceedings of the Nineteenth International Joint Conference on Artificial Intelligence*, pages 947–952, Edinburgh, UK, August 2005.

Susan Epstein. Pragmatic navigation: Reactivity, heuristics, and search. *Artificial Intelligence*, 100(1):275–322, April 1998.

Oren Etzioni. Intelligence without robots (A reply to Brooks). *Artificial Intelligence Magazine*, 14(4), December 1993.

George W. Evans and Seppo Honkapohja. An interview with Thomas J. Sargent. *Macroeconomic Dynamics*, 9:561–583, 2005.

J. Doyne Farmer, Martin Shubik, and Eric Smith. Is economics the next physical science? *Physics Today*, 58(9):37–42, September 2005.

Drew Fudenberg and David K. Levine. *The Theory of Learning in Games*. MIT Press, Cambridge, MA, 1998.

Gerd Gigerenzer and Daniel G. Goldstein. Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review*, 103(4):650–669, 1996.

Gerd Gigerenzer and Reinhard Selten. Rethinking rationality. In Gerd Gigerenzer and Reinhard Selten, editors, *Bounded Rationality: The Adaptive Toolbox*, pages 1–12. The MIT Press, Cambridge, MA, 2001.

J.C. Gittins and D.M. Jones. A dynamic allocation index for the sequential design of experiments. In J. Gani, K. Sakadi, and I. Vinczo, editors, *Progress in Statistics*, pages 241–266. North Holland, Amsterdam, 1974.

Daniel G. Goldstein and Gerd Gigerenzer. Models of ecological rationality: The recognition heuristic. *Psychological Review*, 109(1):75–90, 2002.

E.J. Horvitz. Reasoning about beliefs and actions under computational resource constraints. In *Proceedings of the 3rd AAAI Workshop on Uncer-*

tainty in Artificial Intelligence, pages 429–444, 1987.

Michael Kearns and Satinder Singh. Near-optimal reinforcement learning in polynomial time. *Machine Learning*, 49(2):209–232, 2002.

David Parkes. Bounded rationality. Technical report, University of Pennsylvania, Dept of Computer and Information Sciences, 1999.

Stuart Russell. Rationality and intelligence. *Artificial Intelligence*, 94:57–77, 1997.

Stuart Russell and Peter Norvig. *Artificial Intelligence: A Modern Approach*. Prentice Hall, Upper Saddle River, New Jersey, 2nd edition, 2003.

Herbert A. Simon. A behavioral model of rational choice. *The Quarterly Journal of Economics*, 69:99–118, February 1955.