Proceedings of the Third International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS'04), Columbia University, NY, pp. 930-937.

Universality in Multi-Agent Systems

H. Van Dyke Parunak, Sven Brueckner
Altarum Institute
3520 Green Court, Suite 300
Ann Arbor, MI 48105-1579 USA
{van.parunak, sven.brueckner}@altarum.org

Robert Savit
Dept. of Physics, University of Michigan
Ann Arbor, MI 48104
savit@umich.edu

Abstract

Much research in multi-agent systems reflects the field's origins in classical artificial intelligence, showing how various refinements to the internal reasoning of individual agents improve overall system performance. Sometimes, aspects of a system's behavior are independent of the algorithms used by individual agents. Drawing from an analogy in statistical physics, we term this phenomenon "universality." The underlying concept is that systems whose elements differ widely may nevertheless have common emergent features. We develop a notion of universality in MAS based on the concept's use in its original (physics) setting. We illustrate the concept in several examples, and discuss the implications of MAS universality for the theory and practice of MAS. We speculate that there exists a hierarchy of types of universality. The usual use of the term in statistical mechanics refers to the most refined, simplest, and quantitative, while commonalities among systems that are of interest to the MAS community are associated with somewhat more general and qualitative levels of universality. Such a hierarchy would be an important integrating principle across systems of interacting components including human societies, animal ecologies, multi-agent systems, and atoms and molecules.

1. Introduction

Most research in multi-agent systems (MAS's) reflects the classical AI emphasis on mechanisms for the internal reasoning of individual agents, usually justified by theoretical elegance or simple tests of a few agents in isolation. We have found repeatedly in systems of interacting agents that aspects of the overall system behavior are relatively independent of the algorithms used by individual agents. This paper explores this phenomenon.

One theme of our ongoing research is exploiting statistical physics, a mature science dealing with the emergence of macro-level behaviors from the interactions of micro-level elements, for understanding the relation between individual agent behaviors and overall system behavior. Elsewhere we have developed and exploited parallels between these fields in concepts such as entropy [13], phase shifts [2], and techniques for the analysis of nonlinear data [11], as have other researchers [8]. The empirical phenomenon of system-level insensitivity to detailed agent behavior also has an analogy in statistical physics. Physical phase transitions can be characterized by critical exponents whose values can remain constant over a wide range of different substances, a phenomenon that physicists term "universality."

Universality in physics provides a useful metaphor for understanding how the influence of structural features on emergent system wide behavior can neutralize details of individual behavior, and for guiding further research into the interplay of agent and environment. These insights are important in making engineering decisions concerning MAS's. For instance, more complex agents in general are associated with higher knowledge engineering costs and may require more complex hardware, higher power budgets, and more complex input data to execute than simpler agents. Understanding the environmental circumstances under which a simpler agent can achieve the same functionality will enable more efficient designs.

Section 2 introduces universality in the domain of statistical physics, and Section 3 identifies characteristics that are fruitful for developing the metaphor in multi-agent systems. Section 4 exhibits some examples of universality in MAS that we have observed. Section 5 discusses why universality is important for MAS and suggests directions for further research. Section 6 concludes.

2. Universality in Physics

In physics, universality describes the behavior of the *critical exponents* associated with a *continuous phase transition*. Let us explain the italicized terms.

A phase transition is a mathematical singularity (sometimes a discontinuity) experienced by a system as some parameter varies. A familiar example is the freezing of water at 1 atmosphere and 0° Celsius. In some cases (such as the freezing of water), there is a discontinuity in the density or some analogous quantity generically called an "order parameter" (a monotonic function of the first derivative of a free energy.) In this case, the phase transition is called "discontinuous" or "first order." In other cases , the order parameter itself is continuous, but the rate of change of the order parameter with respect to changes in external variables (such as temperature) is discontinuous. These are "continuous" or "second order" phase transitions.[1] (For the sake of precision, we speak of a "transition" only when we have a formal model that contains a point of nonanalyticity in the limit in which the size of the system approaches infinity. Empirically observed rapid changes in physical quantities unaccompanied by a singularity are called "phase shifts" or "phase changes.")

The parade example of a continuous phase transition occurs in iron at the Curie temperature T_c (760° Celsius). The order parameter is the net magnetization. Below T_c, iron is ferromagnetic: when exposed to a magnetic field, it becomes magnetized in the direction of the field, and this magnetization remains after the field is removed. Above T_c, it becomes paramagnetic: its magnetization is proportional to the applied field, and vanishes when the field is zero. Another example is the critical-point transition of water heated in a closed vessel, where the latent heat vanishes and water and steam cease to be distinct.

Near T_c, all relevant physical quantities (e.g., magnetic susceptibility, specific heat, isothermal compressibility) exhibit power-law behavior. That is, these quantities vary as $\left|T - T_c\right|^{\alpha}$. (In some cases, this behavior is observed approaching Tc from both above and below, in other cases, only from one direction.) The exponent α depends on the physical quantity in question, and may be positive (in which case the quantity vanishes at T_c) or negative (in which case it diverges). Remarkably, the value of the critical exponents are relatively independent of the material being studied. For example, the exponent for the liquid-gas coexistence curve is the same for Ne, A, Kr, Xe, N_2, O_2, CO, and CH_4, although these substances differ widely in atomic weight, molecular structure, and the details of their electrochemical interactions.

Generally, two physical systems will exhibit universality if their interactions have the same spatial dimensionality and interaction symmetry. The symmetry of interactions is captured in the Hamiltonian (or generalized energy) of the overall system, which typically includes terms for potential energy that depend on the interactions of the system elements. The system's symmetry group consists of the set of transformations of the elements under which the value of the Hamiltonian is invariant. For instance, consider a two-dimensional array of spins, each of which can have values of ±1. If the Hamiltonian has only even powers governing interactions among the spins, the spins can be flipped without changing the value of the Hamiltonian. Any two-dimensional system with the same symmetry will have the same critical exponent as this spin system. The aspects of the interaction captured in the symmetry group neutralize the detailed characteristics of the molecules and of their interactions, so that the difference between (say) Ne and CH_4 no longer affects the behavior of the system near the critical point.

[1] One can have transitions in which the rate of change of the order parameter is also continuous, but higher order derivatives may be discontinous. These are also known generically as "continuous" or "higher-order" transitions.

3. Applying the Metaphor

We extend the notion of universality from its simple, precise meaning in statistical mechanics to include any system of interacting elements whose qualitative or quantitiative system-level behavior includes characteristics that are invariant under changes in the individual behavior and detailed interaction of the elements. In these cases we say that differences in individual behavior are *neutralized* by the interactionis We anticipate that a hierarchy of types of universality can be defined, ranging from highly qualitative similarities, through those of an intermediate character, down to the simplest and most quantitative (exemplified by critical exponents in phase transitions). Such a generalization is interesting because it could provide an integrating principle across the field of complex adaptive systems. At this point, we can offer little further formal refinement in definition, but focus on reporting relevant phenomenology, and invite our colleagues to join with us in refining both the definition and our deeper understanding of how these similarities arise.

Universality can be viewed as a manifestation of Ashby's Law of Requisite Variety [1]. As usually stated, this law asserts that the greater the variety of actions in a control system, the greater the variety of disturbances it is able to compensate. A corollary [18] is that the amount of appropriate selection that a controller can perform is limited by the amount of information available. In a physical system, a system's spatial dimensions and the symmetry of the Hamiltonian limit the information to which the molecules can respond, restricting the scope of their actions. Stated in this form, the application to multi-agent systems is straightforward. If agent interactions limit the information available to individuals or their ability to respond nonrandomly to that information, any sophistication above the level appropriate to the environmental information will not make a difference in the system's behavior, and in fact will introduce noise that can limit system efficiency.

Stated in such simple terms, universality seems to be a triviality. Several key characteristics of the phenomenon in physical systems make it non-trivial, and pose analogous challenges in applying the concept to MAS's.

- Universality affects some, but not all, characteristics of a system. In physics, it governs the critical exponents. It does not apply to other physical parameters. For example, the boiling point of a substance is a function of the specific substance. Similarly, an MAS may exhibit universality in some of its system-level characteristics, but not others.
- Universality arises in some, but not all, regimes of a system's operation. (By "regime" we mean a particular set of configuration parameters.) In physics, it arises

near the critical temperature for a phase transition. Similarly a MAS that exhibits universality in one regime may not do so in another. A system's regime determines the nature, structure, and strength of the interaction effects to which individual agents are subjected, and thus the relative importance of individual agent behavior and environment. There are analogs to phase transitions in computational systems [5, 10], including MAS's [19], and the physical analogy suggests that the vicinity of these transitions is a natural place to explore for universality, but our examples show instances of universality that are not near phase shifts.

- Universality defines a set of equivalence classes, not a single class. In physics, one set of substances may share the same critical exponents, but other substances do not. Similarly, we expect that there may be classes of agent architectures within which universality obtains, but between which it does not.

We argue elsewhere [12, 15] that the nonlinearity of agent logic and interactions and the feedbacks among them make MAS's subject to complex and sometimes chaotic behavior. A hallmark of chaos is a system's extreme sensitivity to small differences in initial conditions, a tendency that seems at first glance to contradict our emphasis here that system-level behavior can be universal across different agent behaviors. One distinction between the two phenomena is that sensitivity to initial conditions describes the evolution of a single system through time, while universality compares multiple systems either synchronically or diachronically. For example, the geometry of a system's attractor may exhibit universality, while a specific trajectory on the attractor may be exhibit extreme sensitivity. More generally, the relation between these two phenomena is an opportunity for further study.

The question of which regimes, characteristics, and agent behaviors support universality is at present empirical. Study of this phenomenon must begin with disciplined reports of observed universality, a process that we introduce in this paper. Such reports should clarify the structure and operating regime of the system, identify the aspects of behavior that are observed to be universal, and discuss the possible interaction effects that lead to universality and how differences among individual agent behaviors may be neutralized by or immune to these effects.

4. Examples of Universality in MAS

In this section we present brief reports of a number of manifestations of universality in MAS. These systems were not designed specifically to test for universality. Our observations are post hoc and could be refined by repeating the work with a specific focus on universality. Even in their present form, we believe they offer important insights into how system interactions can neutralize differences among individual agents. Our first report exhibits a very qualitative level of universality between systems that address different but related problems, while the second shows more quantitative universality among different decision protocols in the same problem.

4.1 Regimes in Resource Allocation

Resource allocation is an important application domain for MAS. We have experimented with two abstract models of this domain: the minority game and the distributed graph coloring problem. The decision processes used by individual agents in these systems are quite different from each other (and from realistic resource allocation decision processes), yet both systems exhibit three distinct regimes of behavior that also appear in real-world problems.

Configuration (Minority Game).—In the minority game (MG) [20], a set of N consumers (N odd) repeatedly seek access to $C = N - 1$, resources distributed evenly across $G = 2$ suppliers. In each turn of the game, each consumer selects one of the suppliers. Because $C < N$, one supplier will be overloaded. A consumer receives a point if the supplier it selects turns out not to be overloaded, and nothing if its supplier is overloaded. Each consumer makes its choice with a look-up table, or "strategy," mapping from vectors of m past system states to supplier choice. The larger m, the larger the space of strategies available to consumers, and the more of the past history each consumer can take into account in making its decision. Each consumer has two such tables. After each turn, it computes which strategy would have yielded the greatest payoff throughout the history of the game, and uses that strategy on the next turn. The objective of the game is to maximize the total points awarded to all consumers. The most points that can be awarded in any one turn is $C/2$, which happens when the population of the two suppliers differs by the least possible (one consumer). Thus a useful metric for the system's overall behavior is the variance in the population of either of the suppliers. Because this metric varies inversely with total reward, it measures system inefficiency.

Configuration (Graph Coloring).—Graph coloring is a useful abstraction of many resource allocation problems. Each node in the graph represents a task, each color represents a resource, and an edge between two nodes indicates that a single resource cannot service them simultaneously. Under this model, resource allocation becomes a matter of coloring the graph to minimize the proportion of adjacent nodes that have the same color.

We have been studying [3, 14] the dynamics of a soft real-time graph coloring algorithm developed by the Kestrel Institute to model sensor allocation problems [4]. The graph's dynamics evolve synchronously in discrete time steps. At any time, each node is assigned one of G colors. The assignment of colors may change over time. A node can perceive the current color of its neighbors. A change in a neighbor's color is perceived after a delay of CL ("communication latency") time units. All nodes share a global activation level AL that determines the individual node's probability to activate its local reasoning mechanism at a given time step. If activated, a node re-evaluates its color assignment based on the local metric Degree of Conflict (DoC), computed by dividing the number of neighbors that share the node's color by the overall number of neighbors (*K*). The node calculates the DoC for each of the *G* possible colors, using the perceived color of its neighbors. The node compares the resulting DoC values with the DoC of its current color. Any color whose DoC fulfills a *movement direction* constraint MD is placed into a set of available colors. Two options for MD are "any" (any color is eligible) and "only up" (only colors that would improve the DoC are eligible). The new color of the node is selected from this set of available colors based on a *color selection* rule CS (in this case, pick the color giving the "best," that is, the lowest, DoC).

To measure the performance of this system, we compute the global degree of conflict (GDOC), as the number of edges in the graph that connect two nodes with the same color, times *G*, divided by the overall number of edges. GDOC is 0 for a perfectly colored graph, 1 if all agents make random choices, and greater than 1 if conflicts are worse than under random choice.

Universal Aspects.—Universality manifests itself in the emergence of the same three distinct behavioral regimes in these two very different decision mechanisms.

In the minority game, Figure 1 shows the variation in the system inefficiency (which varies inversely with total reward) as a function of the size of the strategy space (related to *m*). The dashed horizontal line shows the level of inefficiency that would be achieved if all consumers simply picked randomly between the two suppliers at each turn. This Figure shows three distinct regions.

a. For low values of *m*, few distinct strategies are available to the consumers. Many consumers have the same strategy, leading to a thrashing "herding" behavior that is worse than random, but that becomes better than random as *m* increases.
b. As *m* increases to very high values, the space of possible strategies becomes so large that the population of consumers cannot effectively sample it. The population faces an information surfeit compared with its processing resources, and the inefficiency asymptotically approaches the random value.

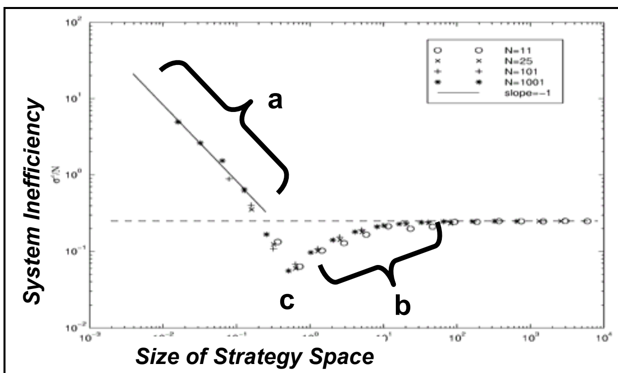

Figure 1: Three regions in the minority game.

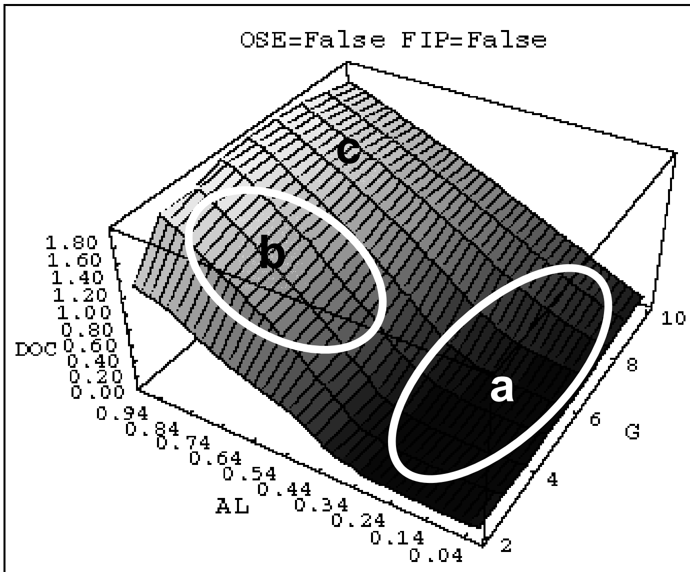

Figure 2: Three regions in graph coloring

c. At an intermediate value, the system faces relatively few choices compared with the amount of available information, and reaches the best solutions.[2]

In the graph coloring algorithm, Figure 2 plots GDOC as a function of the number of colors *G* and the activation level *AL*, for an experiment with $CL = 1$, $N = 60$, $K = 10$. This figure shows three regions that have the same characteristics as those in the minority game.

a. For high activation levels, nodes make decisions with obsolete information. Because of the communication latency, the colors a node ascribes to its neighbors are likely to be incorrect, since they have probably activated and changed color in the meanwhile. Thus a node is likely to choose a color that its neighbors have also chosen, resulting in thrashing and herding.
b. For low activation levels, nodes have accurate information about their neighbors' colors. When many colors are available (high *G*), this information is excessive, since a random choice is likely to be as good as a principled one, and GDOC is near the random value.
c. For low *AL* and low *G*, nodes can make meaningful decisions resulting in better-than-random GDOC. This region is further distinguished from the other two in that quality continues to improve over a longer period of time than in the other two regions.

[2] In our payoff function (the standard one for MG), each consumer on the minority supplier receives one point. One can imagine alternative payoff functions, such as making the award depend on the size of the minority group (for example, a fixed-sum game). The location of this minimum is constant across such alternative payoff functions [9], another example of (quantitative) universality.

Both models exhibit the same three regions: a region of low information with thrashing and herding, a region of excess information resulting in decisions no better than random, and a region where decision capability and available information are roughly balanced, yielding superior performance. These similarities arise even though the underlying decision mechanisms are very different.

Discussion.—The common feature to both the MG decision table mechanism and the graph coloring algorithm is that they are boundedly rational. Any such reasoner will be best suited for a particular amount of information [21]. If the information available is greater than this level, the mechanism will be overwhelmed and give answers essentially no better than random. If it is lower than this level, the mechanism will not be able to break the symmetry among the agents, resulting in herding and thrashing. We hypothesize that these regimes will *not* appear for truly optimizing decision-makers, since an optimal decision by definition uses all of the relevant information available. However, such processes are rarely encountered in practical real-world domains. Real-time constraints impose time limits that effectively turn even an optimizer into a bounded rationalizer when confronted by large enough amounts of information.

4.2 Alternative RAG Decision Algorithms

We have generalized the MG in a number of ways, including more realistic decision rules, to produce a "resource allocation game" (RAG) [16, 19]. Some behavioral options turn out to be neutralized by the structure of resource allocation and do not affect the system's dynamics.

Configuration.—Two important extensions to the MG that make it more realistic include relaxing the constraints that $C = N - 1$ and $G = 2$, and providing a wider range of decision rules on the part of both consumers and suppliers.

Decoupling C and N permits us to explore the dynamics of the system over a range of over- and under-supply situations. We report on the resulting dynamic landscapes elsewhere [16, 19].

In MG, a consumer selects a supplier based on look-up tables ("strategies") driven by the last m outcomes of the game, while a supplier awards a point to all consumers that chose it only if they do not exceed its capacity. Both of these conventions may be arbitrary in some settings, and we explored alternative rules. Our alternatives use the pheromone learning mechanism explained in [14]. Briefly, a decision-maker maintains a vector of pheromone strengths over its options. Each time it chooses an option and experiences a successful outcome, it makes a deposit into the element of the vector corresponding to that option. Over time, all the deposits decay exponentially. To make a decision, the decision-maker spins a roulette wheel weighted by the current pheromone strengths.

On the consumer side, the alternative to strategy tables is to select suppliers based on past success. The decision options are the supplier identities, and success consists of receiving a point. This mechanism differs from the strategy table mechanism in two important ways. First, the scoring process models a consumer's loyalty to suppliers that perform well for it, while the strategy table approach models a consumer's loyalty to a specific decision rule without regard to which supplier is chosen. Second, the strategy table approach uses global knowledge (which supplier was not overloaded in the previous round), while the scoring process uses only local information (whether or not this consumer was successful on the last round).

On the supplier side, MG gives awards only when the supplier is not overloaded. This mechanism ("binary satisfaction") is reasonable for some resources (such as electrical power, in which the resource will die if overloaded so that no consumers are satisfied). In other cases (such as shoppers at a store), all consumers up to the supplier's capacity can be satisfied and the rest turned away empty ("partial satisfaction"). We explored two mechanisms for partial satisfaction. The first allocates the available capacity randomly with equal weight among the consumers who select a supplier. The second uses pheromone learning to model a "frequent buyer" program: the supplier's decision options are the identities of the individual consumers, and success consists of being selected by a consumer.

Our performance metric for MG, based on the standard deviation of supplier populations, is more appropriate for binary than for partial satisfaction. For the RAG, we focus on the mean award rate, which is the probability across the entire population of consumers that a consumer will receive a point on a given attempt. This is a global property of the system, which emerges from the local interactions of the agents.

The result of these options is a set of six possible configurations. Consumers can select suppliers using either strategy tables or supplier scores, while suppliers can award consumers using either binary satisfaction, partial satisfaction with random choice, and partial satisfaction using consumer scores. For each of these configurations, we evaluate the mean award rate as a function of the total capacity of the suppliers. In each plot, the abscissa is the total supply available, ranging from 4 to 24 in steps of two. The ordinate is the mean award rate. Each plot shows both points for 13 individual runs at each level of supply, and a line plotted through the means of the points for each level of supply. The first 1500 time steps of each run are discarded in computing the mean award rate, to allow the systems to stabilize. The demand is set at 14, so for supply below 14 one expects low mean award rates, while for supply above 14 one expects high mean award rates. Figure 3 shows the resulting set of curves.

Universal Aspects.—Figure 3 shows interesting instances of universality at three different levels.

The two approaches to partial satisfaction exhibit full quantitative universality. They are completely indistin-

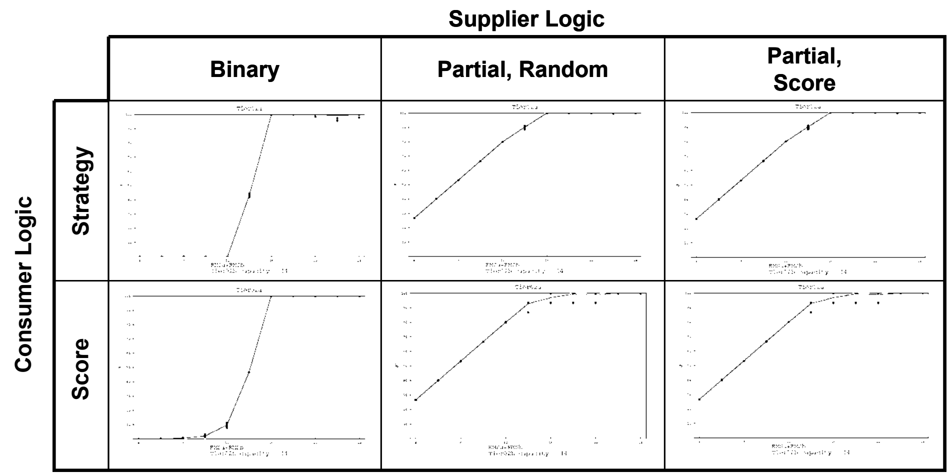

Figure 3: Mean award rate for different configurations for demand (= 14) as a function of supply (ranging from 4 to 24)

guishable. It doesn't matter whether suppliers favor certain customers or treat them all the same, at least from the perspective of the global mean award rate. (It is worth repeating that universality is with respect to some specified system-level characteristic. We do not claim that every aspect of the system is the same in these two cases, only that the mean award rate is.)

The two rows of Figure 3 exhibit universality at a somewhat more qualitative level. The curves have the same general shape and the transitions are in the same places. The only difference is that the change in the mean award rate as the system moves from under- to over-supply is more gradual for score-based consumers than for strategy-based consumers. In terms of the (global) mean award rate, it makes little difference whether consumers select their suppliers based on a history-driven strategy, or whether they prefer specific suppliers.

One might have expected that the configuration in which suppliers learn to prefer certain consumers and consumers learn to prefer certain suppliers (the lower right-hand plot) would yield higher award rates than the others. Since both consumer and supplier preferences are learned over time, this configuration would appear to favor the emergence of specific teams, yielding improved efficiency over the other configurations. But in fact, learned teaming conveys no advantage in terms of mean award rate.

In a still more qualitative sense of universality, the curves for binary and partial satisfaction are alike in that they are all monotonic increasing in supply, though the number of inflection points and the locations of the transitions differ. This broader universality class includes all six configurations.

Discussion.—Consider first the lower level of universality between partial and binary satisfaction. The tendency of both consumer decision processes is to spread the consumers out evenly over the suppliers. If total supply is less than total demand, all suppliers will be oversubscribed. Under binary satisfaction, no points will be awarded, while with partial satisfaction, all available points will be supplied up to the number of consumers. This simple argument leads us to expect at least a quantitative difference in the total amount awarded, and thus in the mean award rate.

Now consider the universality over consumer decision algorithms. The key to maximizing the mean award rate is developing heterogeneous decisions by the consumers so that they are spread out as much as possible across the suppliers. From this perspective, there is no *a priori* reason that either of the two approaches tested (diversification based on strategy or based on preference for a given supplier) should dominate. Both are adaptive algorithms, and both have enough latitude to distribute the consumers over the suppliers. (For the table-based process, this claim depends critically on the depth of history m used in the strategy table, as Figure 1 shows. In these experiments, we use $m = 4$, which is close to the optimal for the size of the system in these experiments.)

The somewhat smoother nature of the transition for score-based decisions reflects the more stochastic nature of these decisions compared with strategy-based decisions. Recall our simple explanation of the non-universality between partial and binary satisfaction. The strategy decision process is deterministic. If the agents do indeed learn to distribute themselves evenly, they will always do so, meaning that none will be successful until supply is at least as large as demand. However, the scoring process is stochastic. An agent who has learned to prefer one supplier will occasionally try the other. When supply is almost adequate, a few defections from one supplier may be enough to enable it to reward the consumers who remain.

The universality over the two approaches to partial satisfaction is not surprising. In either case, a supplier awards as many consumers as it can, so the total number of awards in a given turn of the game does not depend on which selection process is used.

This discussion suggests that in resource allocation problems, system-level constraints such as the ratio of supply to demand are more important than detailed decision processes in determining some system-level outcomes. Individual profit and loss may vary widely depending on individual decisions, but at the system level, the overall structure can neutralize differences in individual reasoning.

5. Why is Universality Important?

The concept of universality is important to MAS researchers and developers for at least three reasons. It provides a rationale for agent-based (and other) modeling, it guides system design and implementation, and it is rich topic for theoretical research.

5.1 Agent-Based Modeling

A major application for MAS is in constructing models of the real world. Such agent-based models (ABM's) are attractive for many reasons [17], not the least because they offer a straightforward mapping to entities in the real-world domain and the interactions among those entities. The results they give are often qualitatively different from (and superior to) those from equation-based models.

In the face of these successes, it is easy to overlook how simplistic the agents in a useful agent-based model can be, compared with the entities in the real world. A parade example is the use of systems such as ISAAC/EINSTein [6] and MANA [7] for modeling land combat. The agents in these systems are extremely simple, based on cellular automata and simple attraction-repulsion rules that determine a soldier's movement. In spite of this simplicity, the models yield emergent patterns of military engagements that are sufficiently realistic to be useful to military planners in evaluating alternative strategies [22]. To an unbiased observer, this success seems almost magical, or at the least coincidental, leading some users to hesitate in trusting the results.

Universality helps explain this unreasonable success. In spite of all the detail that a simple ABM omits, it captures important qualitative features of the interactions among the entities, which in many cases neutralize any detailed refinements in the behavior of individual agents. The existence and widespread manifestation of universality can help build confidence in ABM's, as well as guide in their refinement as users gain experience in how universality manifests itself in specific configurations.

5.2 System Design and Implementation

The designer of any practical system must continually balance cost against benefit. A major driver of cost in MAS is the design of the individual agent. Coarse-grained or BDI agents can support extremely nuanced reasoning, but require detailed (and often costly) knowledge engineering, and typically assume an execution environment comparable to Unix or Windows. Fine-grained architectures inspired by artificial life can be configured by evolution rather than knowledge engineering, and can run on platforms with very limited power and memory requirements, but individual agents have a much more limited behavioral repertoire.

A major decision in a specific multi-agent system is to determine where along this continuum the individual agents should be designed. Understanding universality can facilitate this decision. As we gain experience in the ways that agent interactions constrain the decisions they can make, we can begin to develop design guidelines indicating how much agent sophistication is needed in a given setting. The examples given above suggest that in certain resource allocation problems, very simple agents can collectively do just as good a job as more complex ones, at lower engineering cost.

5.3 Theory

The perspective of universality urges us to shift our research focus from how various agent behaviors change overall system behavior, to the circumstances under which they do not. This shift in focus introduces a host of new theoretical questions that cry out for investigation. Here we enumerate a few of these, roughly ordered to reflect a possible research agenda.

Where does universality occur? Researchers who observe universality in MAS should not brush it off as uninteresting, but document the systems in which it occurs.

How should we describe it? Generalization across examples will lead to a taxonomy of environments or interaction structures and a detailed understanding of the range of decision processes that each of them can neutralize.

How can we measure it? Statistical mechanics measures similarity among individuals in terms of atomic and molecular properties; among interactions in terms of dimensionality and symmetry classes; and among system behaviors in terms of critical exponents. What quantitative metrics would be useful for studying universality in MAS?

What causes it? It seems clear that the driver for universality lies in the interactions among the agents, but what aspects of those interactions are determinative of universality? In statistical mechanics, only dimensionality and symmetry matter for the determination of critical exponents, and differences in detailed electrochemical interaction are irrelevant. Can analogous notions of dimensionality and symmetry be developed for MAS's? Should we focus our attention on the propagation of constraints among agents? Or perhaps on the dynamics of the information flow among them?

How can we predict it? The ultimate objective is a predictive theory that will enable us to tell when differences among agents will lead to system-level differences, and when they will not. Such a theory provides the foundation for addressing in a more disciplined fashion the two benefits discussed in Sections 5.1 and 5.2 above.

We speculate that the similarity we have observed between atomic-level systems and multi-agent systems can be extended to many other systems of interacting components, such as social structures, ecologies, and economies. All these systems consist of individual entities interacting with one another. A refined universality hierarchy could help unify these different disciplines. The MAS community is strongly interdisciplinary, and is well suited to lead research from this new perspective.

6. Conclusion

A common rhetorical pattern in MAS papers is to show how the performance of a proposed innovation differs from some benchmark. Sometimes, it is more instructive to observe when a change in agent algorithms does *not* change the outcome, because it is neutralized by the interactions among the agents. We have observed this effect in multi-agent models of resource allocation, and suspect that it is even more widespread. Understanding the potential for this kind of behavior offers a disciplined way to explain the sometimes surprising success of simple agent-based models, and can point the way to more efficient MAS designs, but its full exploitation requires an ongoing theoretical effort to explain just when universality will occur, how it will manifest itself, and what sorts of individual behaviors are subject to it.

Acknowledgements

This work is supported in part by the DARPA ANTS program, contract F30602-99-C-0202 to Altarum, under Dr. Vijay Raghavan. The views and conclusions in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Defense Advanced Research Projects Agency or the US Government.

References

[1] W. R. Ashby. Requisite variety and its implications for the control of complex systems. *Cybernetica*, 1(2):83-99, 1958.

[2] S. Brueckner and H. V. D. Parunak. Information-Driven Phase Changes in Multi-Agent Coordination. In *Proceedings of Autonomous Agents and Multi-Agent Systems (AAMAS 2003)*, 950-951, 2003.

[3] S. Brueckner and H. V. D. Parunak. Resource-Aware Exploration of Emergent Dynamics of Simulated Systems. In *Proceedings of Autonomous Agents and Multi-Agent Systems (AAMAS 2003)*, 781-788, 2003.

[4] S. Fitzpatrick and L. Meertens. Soft, Real-Time, Distributed Graph Coloring using Decentralized, Synchronous, Stochastic, Iterative-Repair, Anytime Algorithms: A Framework. KES.U.01.5., Kestrel Institute, Palo Alto, CA, 2001. URL ftp://ftp.kestrel.edu/pub/papers/fitzpatrick/eMergeANT-001.pdf.

[5] T. Hogg, B. A. Huberman, and C. Williams. Phase Transitions and the Search Problem. *Artificial Intelligence*, 81:1-15, 1996.

[6] A. Ilachinski. ISAAC/EINSTein: An Artificial-Life Approach to Land Combat. 2001. Web page, http://www.cna.org/isaac/.

[7] M. K. Lauren and R. T. Stephen. Map-Aware Non-uniform Automata (MANA)—A New Zealand Approach to Scenario Modelling. *Journal of Battlefield Technology*, 5(1 (March)):27ff, 2002.

[8] K. Lerman and A. Galstyan. A General Methodology for Mathematical Analysis of Multi-Agent Systems. ISI-TR-529, USC Information Sciences Institute, Marina del Rey, CA, 2001. URL http://www.isi.edu/%7Elerman/papers/isitr529.pdf.

[9] Y. Li, A. VanDeemen, and R. Savit. The Minority Game with Variable Payoffs. *Physica A*, 284(2000):461-477, 2000.

[10] R. Monasson, R. Zecchina, S. Kirkpatrick, B. Selman, and L. Troyansky. Determining computational complexity from characteristic 'phase transitions'. *Nature*, 400(July 8):133-137, 1999.

[11] H. V. D. Parunak. The Heartbeat of the Factory. ERIM, Ann Arbor, 1995. URL http://www.erim.org/~vparunak/heartbt.pdf.

[12] H. V. D. Parunak. From Chaos to Commerce: Practical Issues and Research Opportunities in the Nonlinear Dynamics of Decentralized Manufacturing Systems. In *Proceedings of Second International Workshop on Intelligent Manufacturing Systems*, k15-k25, K.U.Leuven, 1999.

[13] H. V. D. Parunak and S. Brueckner. Entropy and Self-Organization in Multi-Agent Systems. In *Proceedings of The Fifth International Conference on Autonomous Agents (Agents 2001)*, 124-130, ACM, 2001.

[14] H. V. D. Parunak, S. Brueckner, R. Matthews, and J. Sauter. How to Calm Hyperactive Agents. In *Proceedings of Autonomous Agents and Multi-Agent Systems (AAMAS 2003)*, 1092-1093, 2003.

[15] H. V. D. Parunak, S. Brueckner, and J. Sauter. ERIM's Approach to Fine-Grained Agents. In *Proceedings of NASA/JPL Workshop on Radical Agent Concepts (WRAC)*, 2001.

[16] H. V. D. Parunak, S. Brueckner, J. Sauter, and R. Savit. Effort Profiles in Multi-Agent Resource Allocation. In *Proceedings of Autonomous Agents and Multi-Agent Systems (AAMAS02)*, 248-255, 2002.

[17] H. V. D. Parunak, R. Savit, and R. L. Riolo. Agent-Based Modeling vs. Equation-Based Modeling: A Case Study and Users' Guide. In *Proceedings of Multi-agent systems and Agent-based Simulation (MABS'98)*, 10-25, Springer, 1998.

[18] Principia Cybernetica. Law of Requisite Variety. 2003. Web page, http://pespmc1.vub.ac.be/ASC/LAW_VARIE.html.

[19] R. Savit, S. A. Brueckner, H. V. D. Parunak, and J. Sauter. Phase Structure of Resource Allocation Games. *Physics Letters A*, 311:359-364, 2002.

[20] R. Savit, R. Manuca, and R. Riolo. Adaptive Competition, Market Efficiency, and Phase Transitions. *Physical Review Letters*, 82(10):2203-2206, 1999.

[21] T. Takashina and S. Watanabe. The Locality of Information Gathering in Multiagent Systems. In *Proceedings of Second International Conference on Multi-Agent Systems (ICMAS-96)*, 461, 1996.

[22] USMC. Project Albert Home Page. 2003. Web Site, http://www.mcwl.quantico.usmc.mil/divisions/albert/index.asp.