EUROPHYSICS LETTERS

Europhys. Lett., (), pp. ()

Collective Intelligence for Control of Distributed Dynamical Systems

D. H. Wolpert, K. R. Wheeler and K. Tumer

NASA Ames Research Center
Moffett Field, CA 94035

(received ; accepted)

PACS. 07.05.Mh – Neural networks, fuzzy logic, artificial intelligence.
PACS. 89.80.+h – Computer science and technology.
PACS. 89.90.+n – Other areas of general interest to physicists.

Abstract. – We consider the El Farol bar problem, also known as the minority game (W. B. Arthur, *The American Economic Review*, 84(2): 406–411 (1994), D. Challet and Y.C. Zhang, *Physica A*, 256:514 (1998)). We view it as an instance of the general problem of how to configure the nodal elements of a distributed dynamical system so that they do not "work at cross purposes", in that their collective dynamics avoids frustration and thereby achieves a provided global goal. We summarize a mathematical theory for such configuration applicable when (as in the bar problem) the global goal can be expressed as minimizing a global energy function and the nodes can be expressed as minimizers of local free energy functions. We show that a system designed with that theory performs nearly optimally for the bar problem.

Introduction. – In many distributed dynamical systems there is little centralized communication and control among the individual nodal elements. Despite this handicap, typically we wish to design the system so that its dynamical behavior has some desired form. Often the quality of that behavior can be expressed as a (potentially path-dependent) global energy function, G. The associated design problem is particularly interesting when we can also express the individual nodal elements η as minimizers of "local" energy functions γ_η. Given G, this reduces the problem to determining the optimal associated $\{\gamma_\eta\}$.

Because the argument lists of the γ_η may overlap, what action η should take at time t to minimize γ_η may depend on what actions the other nodes take at t. Since without binding contracts η cannot know those other actions ahead of time, it cannot assuredly minimize γ_η in general. We are particularly interested in cases where each η addresses this problem by using machine-learning (ML) techniques to determine its actions. (In its use of such techniques that trade off exploration and exploitation, such an η often approximates a stochastic node following the distribution that minimizes η's free energy — see below.) In such cases the challenge is to choose the $\{\gamma_\eta\}$ so that the associated system of good (but suboptimal) ML-based nodes induces behavior that best minimizes the provided function G.

We refer to a system designed this way, or more generally to a system investigated from this perspective, as a COllective INtelligence (COIN) [14, 15]. To agree with bar problem and game-theory terminology, we refer to the nodes as *agents*, G as (minus) *world utility*, and the $\{\gamma_\eta\}$ as (minus) *private utilities.* As an example of this terminology, a spin glass in which each spin η is at an energy minimum given the states of the other spins is a "Nash equilibrium" of an associated "game" [9], a game formed by identifying each agent with a spin η and its associated private utility function with η's energy function.

Arthur's bar problem [1] can be viewed as a problem in designing COINs. Loosely speaking, in this problem at each time t each agent η decides whether to attend a bar by predicting, based on its previous experience, whether the bar will be too crowded to be "rewarding" at that time, as quantified by a reward function $R_{UD;\eta}$. The greedy nature of the agents frustrates the global goal of maximizing $G = \sum_\eta R_{UD;\eta}$ at t. This is because if most agents think the attendance will be low (and therefore choose to attend), the attendance will actually be high, and vice-versa. This frustration effect makes the bar problem particularly relevant to the study of the physics of emergent behavior in distributed systems [3, 4, 5, 8, 11, 12, 16].

In COIN design we try to avoid such effects by determining new utilities $\{\gamma_\eta\}$ so that all agents trying to minimize those new utilities means that G is also minimized. (Of course, we wish to determine the $\{\gamma_\eta\}$ without first explicitly solving for the minimum of G.) As an analogy, economic systems sometimes have a "tragedy of the commons" (TOC) [10], where each agent's trying maximize its utility results in collective behavior that *minimizes* each agent's utility, and therefore minimizes *minimizes* $G = \sum_\eta R_{UD;\eta}$. One way the TOC is avoided in real-world economies is by reconfiguring the agents' utility functions from $\{R_{UD;\eta}\}$ to a set of $\{\gamma_\eta\}$ that results in better G, for example via punitive legislation like anti-trust regulations. Such utility modification is exactly the approach used in COIN design.

We recently applied such COIN design to network packet routing [15]. In conventional packet routing each router uses a myopic shortest path algorithm (SPA), with no concern for side-effects of its decisions on an external world utility like global throughput (e.g., for whether those decisions induce bottlenecks). We found that a COIN-based system has significantly better throughput than does a conventional SPA [15], even when the agents in that system had to predict quantities (e.g., delays on links) that were directly provided to the SPA.

In this paper we confront frustration effects more directly, in the context of the bar problem. In the next section we present (a small portion of) the theory of COINs. Then we present experiments applying that theory to the distributed control of the agents in the bar problem. Those experiments indicate that by using COIN theory we can avoid the frustration in the bar problem and thereby achieve almost perfect minimization of the global energy.

Theory of COINs. – We consider the state of the system across a set of consecutive time steps, $t \in \{0, 1, ...\}$. Without loss of generality, all relevant characteristics of agent η at time t — including its internal parameters at that time as well as its externally visible actions — are encapsulated by a Euclidean vector $\underline{\zeta}_{\eta,t}$, the *state* of agent η at time t. $\underline{\zeta}_{,t}$ is the set of the states of all agents at t, and $\underline{\zeta}$ is the state of all agents across all time.

So **world utility** is $G(\underline{\zeta})$, and when η is an ML algorithm "striving to increase" its **private utility**, we write that utility as $\gamma_\eta(\underline{\zeta})$. The mathematics is generalized beyond such ML-based agents through an artificial construct: the **personal utilities** $\{g_\eta(\underline{\zeta})\}$. We restrict attention to utilities of the form $\sum_t R_t(\underline{\zeta}_{,t})$ for **reward functions** R_t.

We are interested in systems whose dynamics is deterministic. (This covers in particular any system run on a digital computer.) We indicate that dynamics by writing $\underline{\zeta} = C(\underline{\zeta}_{,0})$. So all characteristics of an agent η at $t = 0$ that affects the ensuing dynamics of the system, including in particular its private utility if it has one, must be included in $\underline{\zeta}_{\eta,0}$.

Definition: A system is **factored** if for each agent η individually,

$$g_\eta(C(\underline{\zeta}_{,0})) \geq g_\eta(C(\underline{\zeta}'_{,0})) \quad \Leftrightarrow \quad G(C(\underline{\zeta}_{,0})) \geq G(C(\underline{\zeta}'_{,0})) \ , \tag{1}$$

for all pairs $\underline{\zeta}_{,0}$ and $\underline{\zeta}'_{,0}$ that differ only for node η.

For a factored system, the side effects of a change to η's $t = 0$ state that increases its personal utility cannot decrease world utility. If the separate agents have high personal utilities, by luck or by design, then they have not frustrated each other, as far as G is concerned.

The definition of factored is carefully crafted. In particular, it does *not* concern changes in the value of the utility of agents other than the one whose state is varied. Nor does it concern changes to the states of more than one agent at once. Indeed, consider the following alternative desideratum to having the system be factored: any change to $\underline{\zeta}_{,0}$ that simultaneously improves all agents' ensuing utilities must also improve world utility. Although it seems quite reasonable, there are systems that obey this desideratum and yet quickly evolve to a *minimum* of world utility. For example, any system that has $G(\underline{\zeta}) = \sum_\eta g_\eta(\underline{\zeta})$ obeys this desideratum, and yet as shown below, such systems entail a TOC in the the bar problem.

For a factored system, when every agents' personal utility is optimizal, given the other agents' behavior, world utility is at a critical point [14]. In game-theoretic terms, optimal global behavior corresponds to the agents' reaching a personal utility Nash equilibrium for such systems [9]. Accordingly, there can be no TOC for a factored system.

As a trivial example, if $g_\eta = G \ \forall \eta$, then the system is factored, regardless of C. However there exist other, often preferable sets of $\{g_\eta\}$, as illustrated in the following development.

Definition: The $(t = 0)$ **effect set** of node η at $\underline{\zeta}$, $C^{eff}_\eta(\underline{\zeta})$, is the set of all components $\underline{\zeta}_{\eta',t'}$ for which $\partial_{\underline{\zeta}_{\eta,0}}(C(\underline{\zeta}_{,0}))_{\eta',t'} \neq \vec{0}$. C^{eff}_η with no specification of $\underline{\zeta}$ is defined as $\cup_{\underline{\zeta} \in C} C^{eff}_\eta(\underline{\zeta})$.

Definition: Let σ be a set of agent-time pairs. $\mathrm{CL}_\eta(\underline{\zeta})$ is $\underline{\zeta}$ modified by "clamping" the states corresponding to all elements of σ to some arbitrary pre-fixed value, here taken to be $\vec{0}$. The **wonderful life utility** (WLU) for σ at $\underline{\zeta}$ is defined as:

$$WLU_\sigma(\underline{\zeta}) \equiv G(\underline{\zeta}) - G(\mathrm{CL}_\sigma(\underline{\zeta})) \ . \tag{2}$$

In particular, the WLU for the effect set of node η is $G(\underline{\zeta}) - G(\mathrm{CL}_{C^{eff}_\eta}(\underline{\zeta}))$.

η's effect set WLU is analogous to the change world utility would undergo had node η "never existed". (Hence the name of this utility - cf. the Frank Capra movie.) However CL(.) is a purely "fictional", counter-factual mapping, in that it produces a new $\underline{\zeta}$ without taking into account the system's dynamics. The sequence of states produced by the clamping operation in the definition of the WLU need not be consistent with the dynamical laws embodied in C. This is a crucial strength of effect set WLU. It means that to evaluate that WLU we do *not* try to infer how the system would have evolved if node η's state were set to $\vec{0}$ at time 0 and the system re-evolved. So long as we know G and the full $\underline{\zeta}$, and can accurately estimate what agent-time pairs comprise C^{eff}_η, we know the value of η's effect set WLU — even if we know nothing of the details of the dynamics of the system.

Theorem 1: A COIN is factored if $g_\eta = WLU_{C^{eff}_\eta} \ \forall \eta$ (proof in [14]).

If our system is factored with respect to personal utilities $\{g_\eta\}$, then we want each $\underline{\zeta}_{\eta,0}$ to be a state with as high a value of $g_\eta(C(\underline{\zeta}_{,0}))$ as possible. Assuming η is ML-based and able to achieve close to the largest possible value of any private utility specified in $\underline{\zeta}_{\eta,0}$, we would likely be in such a state of high personal utility if η's private utility were set to the associated personal utility: $\gamma_\eta \equiv \underline{\zeta}_{\eta,0;\mathrm{private-utility}} = g_\eta$. Enforcing this equality, our problem becomes determining what $\{\gamma_\eta\}$ the agents will best be able to maximize while also causing dynamics that is factored with respect to the $\{\gamma_\eta\}$.

Now regardless of $C(.)$, both $\gamma_\eta = G\ \forall\eta$ and $\gamma_\eta = WLU_{C^{eff}_\eta}\ \forall\eta$ are factored systems (for $g_\eta = \gamma_\eta$). However since each agent is operating in a large system, it may experience difficulty discerning the effects of its actions on G when G sensitively depends on all components of the system. Therefore each η may have difficulty learning how to achieve high γ_η when $\gamma_\eta = G$. This problem can be obviated using effect set WLU as the private utility, since the subtraction of the clamped term removes some of the "noise" of the activity of other agents, leaving only the underlying "signal" of how the agent in question affects the utility.

We can quantify this signal/noise effect by comparing the ramifications on the private utilities arising from changes to $\underline{\zeta}_{\eta,0}$ with the ramifications arising from changes to $\underline{\zeta}_{\hat{}\eta,0}$, where $\hat{}\eta$ represents all nodes *other* than η. We call this quantification the **learnability** $\lambda_{\eta,\gamma_\eta}(\underline{\zeta})$:

$$\lambda_{\eta,\gamma_\eta}(\underline{\zeta}) \equiv \frac{\|\vec{\nabla}_{\underline{\zeta}_{\eta,0}} \gamma_\eta(C(\underline{\zeta}_{,0}))\|}{\|\vec{\nabla}_{\underline{\zeta}_{\hat{}\eta,0}} \gamma_\eta(C(\underline{\zeta}_{,0}))\|} \ . \tag{3}$$

Theorem 2: Let σ be a set containing C^{eff}_η. Then

$$\frac{\lambda_{\eta,WLU_\sigma}(\underline{\zeta})}{\lambda_{\eta,G}(\underline{\zeta})} = \frac{\|\vec{\nabla}_{\underline{\zeta}_{\hat{}\eta,0}} G(C(\underline{\zeta}_{,0}))}{\|\vec{\nabla}_{\underline{\zeta}_{\hat{}\eta,0}} G(C(\underline{\zeta}_{,0})) - \vec{\nabla}_{\underline{\zeta}_{\hat{}\eta,0}} G(\mathrm{CL}_\sigma(C(\underline{\zeta}_{,0})))\|} \text{proof in [14].}$$

This ratio of gradients should be large whenever σ is a small part of the system, so that the clamping won't affect G's dependence on $\underline{\zeta}_{\hat{}\eta,0}$ much, and therefore that dependence will approximately cancel in the denominator term. In such cases, WLU will be factored just as G is, but far more learnable. The experiments presented below illustrate the power of this fact in the context of the bar problem, where one can readily approximate effect set WLU and therefore use a utility for which the conditions in Thm.'s 1 and 2 should approximately hold.

Experiments. – We modified Arthur's original problem to be more general, and since we are not interested here in directly comparing our results to those in [1, 2, 4, 8], we use a more conventional ML algorithm than the ones investigated in [1, 2, 4, 8, 13], an algorithm that approximately minimizes free energy. These modifications are similar to those in [3].

There are N agents, each picking one of seven nights to attend a bar the following week, a process that is then repeated. In each week, each agent's pick is determined by its predictions of the associated rewards it would receive. Each such prediction in turn is based solely upon the rewards received by the agent in those preceding weeks in which it made that pick.

The world utility is $G(\underline{\zeta}) = \sum_t R_G(\underline{\zeta}_{,t})$, where $R_G(\underline{\zeta}_{,t}) \equiv \sum_{k=1}^{7} \phi_k(x_k(\underline{\zeta},t))$, $x_k(\underline{\zeta},t)$ is the total attendance on night k at week t, $\phi_k(y) \equiv \alpha_k y \exp{(-y/c)}$; and c and the $\{\alpha_k\}$ are real-valued parameters. Intuitively, this G is the sum of the "world rewards" for each night in each week. Our choice of $\phi_k(.)$ means that when too few agents attend some night in some week, the bar suffers from lack of activity and therefore the world reward is low. Conversely, when there are too many agents the bar is overcrowded and the reward is again low.

Two different $\vec{\alpha}$'s are investigated. One treats all nights equally; $\vec{\alpha} = [1\ 1\ 1\ 1\ 1\ 1\ 1]$. The other is only concerned with one night; $\vec{\alpha} = [0\ 0\ 0\ 7\ 0\ 0\ 0]$. $c = 6$ and N is 4 times the number of agents needed to allow c agents to attend the bar on each of the seven nights, i.e., there are $4 \times 6 \times 7 = 168$ agents. For the purposes of the CL operation, an agent's action at time t is represented as a unary seven-dimensional vector, so the "clamped pick" is (0,0,0,0,0,0,0).

Each η has a 7-dimensional vector representing its estimate of the reward it would receive for attending each night of the week. At the end of each week, the component of this vector corresponding to the night just attended is proportionally adjusted towards the actual reward just received. At the beginning of the succeeding week, to trade off exploration and exploitation, η picks the night to attend randomly using a Boltzmann distribution with 7 energies $\epsilon_i(\eta)$ given by the components of η's estimated rewards vector, and with a temperature decaying in time. This distribution of course minimizes the expected free energy of η, $E(\epsilon(\eta))-TS$, or equivalanetly maximizes entropy S subject to having expected energy given by T. This learning algorithm is similar to Claus and Boutilier's independent learner algorithm [6].

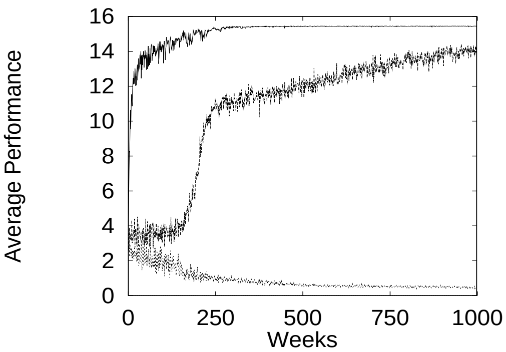
We considered three agent reward functions, using the same learning parameters (learning rate, Boltzmann temperature, decay rates, etc.) for each. The first reward function had $\gamma_\eta = G \ \forall \eta$, i.e., agent η's reward function equals R_G. The other two reward functions are:

$$\begin{array}{lll} \text{Uniform Division (UD):} & R_{UD;\eta}(\underline{\zeta}_{,t}) & \equiv \phi_{d_\eta}(x_{d_\eta}(\underline{\zeta},t))/x_{d_\eta}(\underline{\zeta},t) \\ \text{Wonderful Life (WL):} & R_{WL;\eta}(\underline{\zeta}_{,t}) & \equiv R_G(\underline{\zeta}_{,t}) - R_G(CL_\eta(\underline{\zeta}_{,t})) \ , \end{array}$$

where d_η is the night picked by η. The original version of the bar problem in the physics literature [4] is a special case where there are two "nights" in the week (one of which corresponds to "staying at home"); $\vec{\alpha}$ is uniform; $\phi_k(x_k) = min_i(x_i)\delta_{k,argmin_i(x_i)}$; and $R_{UD\eta}$ is used.

The conventional R_{UD} reward is a "natural" reward function to use; each night's total reward is uniformly divided among the agents attending that night. In particular, if $g_\eta = \gamma_\eta \equiv \sum_t R_{UD;\eta}(\underline{\zeta},t)$, $G(\underline{\zeta}) = \sum_\eta g_\eta(\underline{\zeta})$, so the "alternative desideratum" discussed above is met. In contrast, R_G results in the system meeting the desideratum of factoredness. R_G suffers from poor learnability, at least in comparison to that of R_{WL}; by Eq. 3 the ratio of learnabilities is approximately 11 (see [14] for details). As another point of comparison, to evaluate R_{WL} each agent only needs to know the total attendance on the night it attended, unlike with R_G, which requires centralized communication concerning all 7 nights.

Finally, in the bar problem the only interaction between any pair of agents is indirect, via small effects on each others' rewards; each η's action at time t has its primary effect on η's own future actions. So the effect set of η's entire sequence of actions is well-approximated by $\underline{\zeta}_{\eta,}$. In turn, since that sequence is all that is directly affected by the choice of η's private utility, the effect set of $\underline{\zeta}_{\eta,0;\text{private}-\text{utility}}$ can be approximated by $\underline{\zeta}_{\eta,}$, and therefore so can the effect set of the full $\underline{\zeta}_{\eta,0}$. Therefore we can approximate the effect set WLU for η as $\sum_t R_{WLU;\eta}(\underline{\zeta}_{\eta,t;\text{pick}})$. So we expect that use of $R_{WLU;\eta}$ should result in (close to) factored dynamics.



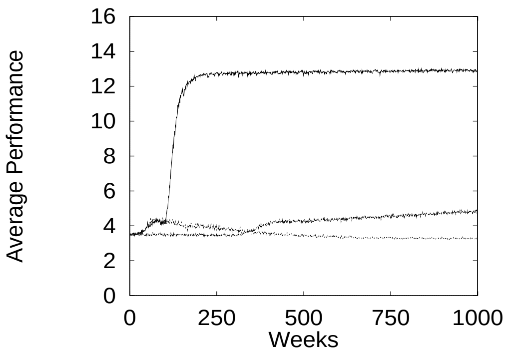


Fig. 1. – Average world reward when $\vec{\alpha} = [0\ 0\ 0\ 7\ 0\ 0\ 0]$ (left) and when $\vec{\alpha} = [1\ 1\ 1\ 1\ 1\ 1\ 1]$ (right). In both plots the top curve is R_{WL}, middle is R_G, and bottom is R_{UD}.

Figure 1 graphs world reward value as a function of time, averaged over 50 runs, for all three reward functions, for both $\vec{\alpha} = [1\ 1\ 1\ 1\ 1\ 1\ 1]$ and $\vec{\alpha} = [0\ 0\ 0\ 7\ 0\ 0\ 0]$. Performance with R_G eventually converges to the global optimum. This agrees with the results obtained by Crites [7] for the bank of elevators control problem. Systems using R_{WL} also converged to optimal performance. This indicates that in the bar problem γ_η's effect set is sufficiently well-approximated by η's future actions so that the conclusions of theorems 1 and 2 hold.

However since the R_{WL} reward has better "signal to noise" than than the R_G reward (see above), convergence with R_{WL} is far quicker than with R_G. Indeed, when $\vec{\alpha} = [0\ 0\ 0\ 7\ 0\ 0\ 0]$, systems using R_G converge in 1250 weeks, which is 5 times worse than the systems using R_{WL}. When $\vec{\alpha} = [1\ 1\ 1\ 1\ 1\ 1\ 1]$ systems take 6500 weeks to converge with R_G, which is more than *30 times* worse than the time with R_{WL}. This slow convergence of systems using R_G is a result of the reward signal being "diluted" by the large number of agents in the system.

In contrast to the behavior for COIN theory-based reward functions, use of conventional R_{UD} reward results in very poor world reward values that deteriorated with time. This is an instance of the TOC. For example, when $\vec{\alpha} = [0\ 0\ 0\ 7\ 0\ 0\ 0]$, it is in every agent's interest to attend the same night — but their doing so shrinks the world reward "pie" that must be divided among all agents. A similar TOC occurs when $\vec{\alpha}$ is uniform. This is illustrated in fig. 2 which shows a typical example of $\{x_k(\underline{\zeta}, t)\}$ for each of the three reward functions for $t = 2000$. In this example using R_{WL} results in optimal performance, with 6 agents each on 6 separate nights, and the remaining 132 agents on one night (average world reward of 13.05). In contrast, R_{UD} results in a uniform distribution of agents and has the lowest average world reward (3.25). Use pf R_G results in an intermediate average world reward (6.01).

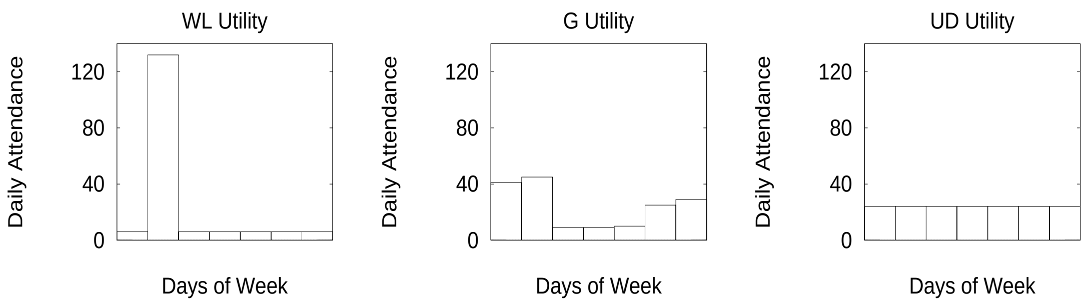


Fig. 2. – Typical daily attendance when $\vec{\alpha} = [1\ 1\ 1\ 1\ 1\ 1\ 1]$ for R_{WL}, R_G, and R_{UD}, respectively.

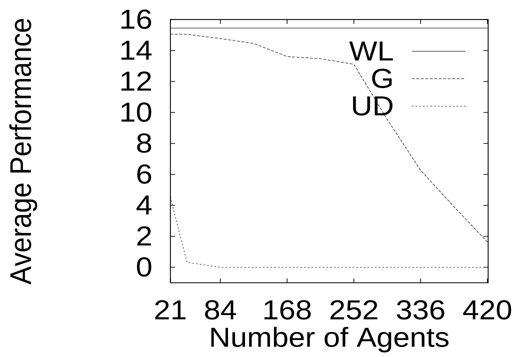


Fig. 3. – Behavior of each reward function with respect to the number of agents for $\vec{\alpha} = [0\ 0\ 0\ 7\ 0\ 0\ 0]$.

Figure 3 shows how performance at $t = 2000$ scales with N for each reward function for $\vec{\alpha} = [0\ 0\ 0\ 7\ 0\ 0\ 0]$. Systems using R_{UD} perform poorly regardless of N. Systems using R_G

perform well when N is low. As N increases however, it becomes increasingly difficult for the agents to extract the information they need from R_G. (This problem is significantly worse for uniform $\vec{\alpha}$.) Systems using R_{WL} overcome this learnability problem because R_{WL} is based on clamping of all agents but one, and therefore is not appreciably affected by N.

Conclusion. – The theory of COINs is concerned with distributed systems of controllers in which each controller strives to minimize an associated local energy function. That theory suggest how to initialize and then update those local energy functions so that the resultant global dynamics will achieve a global goal. In this paper we present a summary of the part of that theory dealing with how to initialize the local energy functions. We present experiments applying that theory to the control of individual agents in difficult variants of Arthur's El Farol Bar problem. In those experiments, the COINs quickly achieve nearly optimal performance, in contrast to the other systems we investigated. This demonstrates that even when the conditions required by the initialization theorems of COIN theory do not hold exactly, they often hold well enough so that they can be applied with confidence. In particular the COINs automatically avoid the tragedy of the commons inherent in the bar problem.

REFERENCES

[1] W. B. Arthur. Complexity in economic theory: Inductive reasoning and bounded rationality. *The American Economic Review*, 84(2):406–411, May 1994.

[2] G. Caldarelli, M. Marsili, and Y. C. Zhang. A prototype model of stock exchange. *Europhys. Letters*, 40:479–484, 1997.

[3] A. Cavagna, Garrahan J. P., I. Giardina, and D. Sherrington. A thermal model for adaptive competition in a market. preprint cond-mat/9903415 v.3, July 1999, and references therein.

[4] D. Challet and Y. C. Zhang. On the minority game: Analytical and numerical studies. *Physica A*, 256:514, 1998, and references therein.

[5] J. Cheng. The mixed strategy equilibria and adaptive dynamics in the bar problem. Technical report, Santa Fe Institute Computational Economics Workshop, 1997.

[6] C. Claus and C. Boutilier. The dynamics of reinforcement learning cooperative multiagent systems. In *Proceedings of the Fifteenth National Conference on Artificial Intelligence*, pages 746–752, June 1998.

[7] R. H. Crites and A. G. Barto. Improving elevator performance using reinforcement learning. In D. S. Touretzky, M. C. Mozer, and M. E. Hasselmo, editors, *Advances in Neural Information Processing Systems - 8*, pages 1017–1023. MIT Press, 1996.

[8] M. A. R. de Cara, O. Pla, and F. Guinea. Competition, efficiency and collective behavior in the "El Farol" bar model. preprint cond-mat/9811162 (to appear in European Physics Journal B), November 1998.

[9] D. Fudenberg and J. Tirole. *Game Theory.* MIT Press, Cambridge, MA, 1991.

[10] G. Hardin. The tragedy of the commons. *Science*, 162:1243–1248, 1968.

[11] N. F. Johnson, S. Jarvis, R. Jonson, P. Cheung, Y. R. Kwong, and P. M. Hui. Volatility and agent adaptability in a self-organizing market. preprint cond-mat/9802177, February 1998.

[12] R. Savit, R. Manuca, and R. Riolo. Adaptive competition, market efficiency, phase transitions and spin-glasses. preprint cond-mat/9712006, December 1997.

[13] W. A. Sethares and A. M. Bell. An adaptive solution to the El Farol problem. In *Proceedings. of the Thirty-Sixth Annual Allerton Conference on Communication, Control, and Computing*, Allerton, IL, 1998. (Invited).

[14] D. H. Wolpert and K. Tumer. An Introduction to Collective Intelligence. In J. M. Bradshaw, editor, *Handbook of Agent technology.* AAAI Press/MIT Press, 1999. to appear.

[15] D. H. Wolpert, K. Tumer, and J. Frank. Using collective intelligence to route internet traffic. In *Advances in Neural Information Processing Systems - 11.* MIT Press, 1999.

[16] Y. C. Zhang. Modeling market mechanism with evolutionary games. preprint cond-mat/9803308, 1998.