Simulating Network Behavioral Dynamics by using a Multi-agent approach driven by ACT-R Cognitive Architecture

Oscar J. Romero
Christian Lebiere
Carnegie Mellon University
5000 Forbes Avenue
Pittsburgh, PA 15213
oscarr@andrew.cmu.edu, cl@cmu.edu

Keywords:
Cognitive Modeling, Social Behavior, Network Science, multi-agent systems

ABSTRACT: *Network Science has reported a considerable amount of human-subject experiments on which individuals have to carry out different kind of coordination games such as coloring and consensus problems in order to observe the behavioral dynamics behind the decision-making process. We have focused on the experiments carried out by Kearns (Kearns, 2010) on which were found strong correlations between the influence and other features of individual and social behavior during the execution of both kind of experiments (the coloring and consensus problems). The aim of this paper is to identify the computational underpinnings of social network behavioral through computational agent modeling using a constrained cognitive architecture framework. Final results demonstrate a strong similarity between human and model dynamics that reflects how complex network behavior can emerge from simple cognitive agents.*

1. Introduction

A significantly amount of studies in the field of network science focus on the analysis of coordination games which are performed in a networked and decentralized fashion. Whereas coordination and cooperation games such as the Prisoners Dilemma and other games have been extensively studied with human subjects over the years (Colin, 2003), (Russel, 1996) , (Reinhard, 1986}, behavioral studies of coordination on networks are more recent. In (Kearns, 2009), study coloring and related problems on networks, although they do not focus on a particular parameterized family of networks as it is done in (Kearns, 2010). McCubbins (McCubbins, 2009) and Kearns (Kearns, 2006) both observe that adding connections makes the coloring problem easier.

The main goal pursued in (Kearns, 2010) is to analyze the dynamics of the process by which players reach coordinated choice and the role of networks in this process. Due to the main objective of our work is to design a cognitive model of the experiments done by (Kearns, 2010), we focused on both the methodological aspects and the behavioral results obtained in each human-subject experiment in order to recreate these behaviors through cognitive models that were then executed in the ACT- R cognitive architecture.

2. Experimental Methodology

The work developed here is based on a line of research at the University of Pennsylvania in controlled human-subject experiments on strategic behavior in social networks (Kearns, 2006), (Kearns, 2009), (Kearns, 2010). In the following, the experiment settings as performed in (Kearns, 2010) are described. First of all, two kind of experiments were performed: in the coloring problem individuals had to choose a color which were different from any of its neighbors and in the consensus problem on which individuals had to reach an unanimous consensus while receiving opposing incentives. The network is a chain of six cliques (complete sub-networks) of six vertexes each, for a total of 36 vertexes, as shown in Fig. 1.

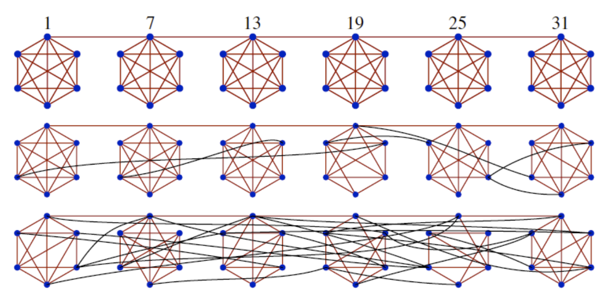

Figure1. Three sample networks used in the experiments. The top one is the baseline network, being a chain of cliques with q = 0, from which all other networks were derived by random edge rewiring. The second network had q = 0.1, and the third had q = 0.2. The six numbered vertices are called "connectors", and the five edges connecting them were retained in all networks. Taken from (Kearns, 2010).

Additionally, there is a probability q ∈ [0,1]. For any fixed value of q, each edge not connecting two cliques in the baseline network is independently "rewired" with probability q. We used values of q ∈ {0;0.1;0.2;0.4;0.6;0.8;1}. For the consensus experiment, nine colors were arbitrarily allowed whereas for the coloring experiment there was a well-defined minimum of colors required for a solution to exist (the so-called chromatic number of each network).

In the consensus experiments each subject received two dollars if a global (unanimous) consensus to any single color was reached, and zero dollars otherwise. In the coloring experiments subjects received two dollars if a valid global coloring was reached, and zero dollars otherwise. Each player was given only a partial or *local neighborhood* view of the network (only the inter-connected nodes are seen). Finally, both experiments had opposite network dynamics.

3. Cognitive Architecture

The cognitive model was developed using the ACT-R cognitive architecture (Anderson, 1998), (Anderson, 2004). Cognitive architectures are computational representations of invariant cognitive mechanisms specified by unified theories of cognition. ACT-R is a modular architecture, reflecting neural constraints, composed of asynchronous modules coordinated through a central procedural system as depicted in figure 2.

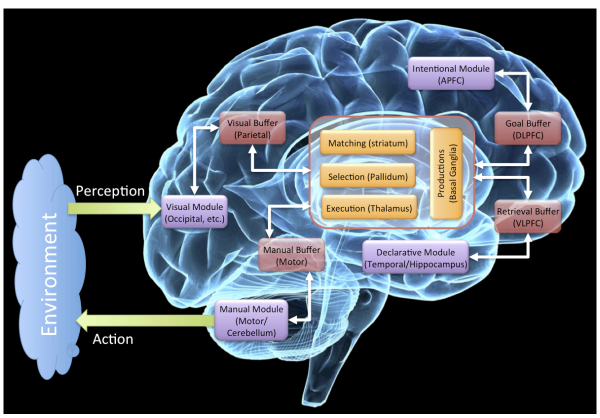

Figure 2. ACT-R Cognitive Architecture

The procedural system is in charge of behavior selection and more generally the synchronization of the flow of information between the other modules. It is implemented as a production system where competing production rules are selected based on their utilities, learning through a reinforcement mechanism from the rewards and costs associated with their actions. The production system conditions are matched against limited-capacity buffers that control the interaction with the other modules by enabling a single command (e.g., retrieval of information, focus of visual attention) to be given at a time to a given module, and a single result to be returned (e.g., chunk retrieved from memory, visual item encoded). A declarative memory module holds both short-term information, such as the details of the current situation, as well as long-term knowledge, such as the procedural rules to follow. Access to memory is controlled by an activation calculus that determines the availability of chunks of information according to their history of use such as recency, frequency, and degree of semantic match. Learning mechanisms control both the automatic acquisition of symbolic structures such as production rules and declarative chunks, and the tuning of their subsymbolic parameters (utility and activation) to the structure of the environment. The perceptual-motor modules reflect human factor limitations such as attentional bottlenecks. Individual differences can be represented both in terms of differences in procedural skills and declarative knowledge, as well as in terms of architectural parameters controlling basic cognitive processes such as spreading of activation.

4. Cognitive Modeling

In the following, the computational cognitive mechanisms used for simulating the results of the social experiments obtained in (Kearns, 2010) will be described.

4.1 Symbolic Productions

As mentioned before, the procedural system uses production rules, which interact with different kind of buffers (retrieval, imaginal, declarative, visual, motor, and others) in order to carry the reasoning and inference process out according to the information that every node of the network senses from its environment and acts over it. We have implemented several strategies to model the social experiments, and every one of these has a set of different productions as described below:

Productions for the Consensus problem: some opposing productions compete against the others to obtain the global control over the decision-making process that performs the agent. Some productions follow the majority, some others follow that agent which is the most influential over the neighborhood[1], as shown in in fig. 3; some other productions just keep the same color no matter if the environmental conditions are not favorable for that, that is, the stubborn productions. ACT-R productions are written in an enriched language that not only allows accessing and manipulating buffer contents but also allows doing as many complex validations as necessary and triggering actions.

1 That is, that agent which has more unseen connections, keeps its color for more time and when it changes its color there are a considerable amount of seen connected agents that do the same

```
(p increasing-then-change-to-majority
   =visual-location>
     isa visual-location
     !eval!   (eq =increasing true)
==>
   +vocal>
     isa speak
     string "change-to-majority")
```

a) If the dominant color is increasing then change to the dominant color

```
(p decreasing-same-majority-then-change-influential
   =visual-location>
     isa visual-location
     !eval!   (eq =decreasing true)
     !eval!   (eq =current-color =majority-color)
==>
   +vocal>
     isa speak
     string "change-to-influential")
```

b) if current color is the same as the majority and its amount is decreasing then change to the most influential agent

```
(p increasing-then-keep-color
   =visual-location>
     isa visual-location
     !eval!   (eq =increasing true)
==>
   +vocal>
     isa speak
     string "keep-color")
```

c) If global consensus is increasing then keep the same color

```
(p decreasing-different-majority-then-change-majority
   =visual-location>
     isa visual-location
     !eval!   (eq =decreasing true)
     !eval!   (not (eq =current-color =majority-color))
==>
   +vocal>
     isa speak
     string "change-to-majority")
```

d) If current color is not the same as the majority and its amount is decreasing then change to majority

Figure 3. Consensus: fragment of productions related to the color changes in the cognitive agent's neighborhood

Productions for the Coloring problem: The basic idea behind this cognitive modeling is to seek an strategy based on both past decisions and current situation that helps the agent to avoid conflicting coloring connections. Stubborn and wrong productions were modeled as well but not presented in order to keep the simplicity.

As you can infer from figure 3, 4, and 5 all the productions generate opposing tensions and a continuous competence for being the production to be fired. For example, "increasing-then-change-majority" production senses the same information as "increasing-then-keep-color" production but they trigger different actions: the first one will change its color in order to follow the majority and the second one will keep its current color if the majority is increasing (whichever its color is). Similar antagonisms are observed in the rest of productions. Due to the fact that multiple productions may match the same sensory input or that sometimes there is not a production, which has a perfect match with the sensory input, a selection process based on production utilities and partial matching is required.

```
(p stuck-then-change-to-influential
   =visual-location>
     isa visual-location
     !eval!   (eq =increasing false)
==>
   +vocal>
     isa speak
     string "change-to-influential")
```

a) If global consensus is stuck then choose statistically the color of an influential agent

```
(p stuck-then-change-another-majority
   =visual-location>
     isa visual-location
     !eval!   (eq =increasing false)
   ==>
   +vocal>
     isa speak
     string "change-another-majority")
```

b) If global consensus is stuck then choose statistically the color of another majority node

```
(p stable
   =visual-location>
     isa visual-location
     !eval!   (eq =increasing true)
==>
   +vocal>
     isa speak
     string "keep-color")
```

c) If the global consensus is increasing then keeps the same current color

```
(p stubborn
   =visual-location>
     isa visual-location
     !eval!   (eq =decreasing true)
     !eval!   (not (eq =current-color =majority-color))
==>
   +vocal>
     isa speak
     string "keep-color")
```

d) If current color is not the same as majority and it is decreasing then keeps the same color.

Figure 4. Consensus: Fragment of productions related to the cognitive agent's internal motivations

It is important to remark that both experiments were running each one by using three different decision-making strategies: 1) a purely deterministic process on which the highest values were always selected (e.g., the majority or the most influential agent for the consensus problem and the *change-to-another-color* action for the coloring problem); 2) an stochastic selection process based on the Boltzman equation (Anderson, 2004) as shown in equation 1; and 3) a second version of the stochastic process using a gradually decreasing value for the temperature, similar to the simulated annealing approach (Kirkpatrick, 1983).

$$P_i = \frac{e^{M_j/t}}{\sum_j e^{M_j/t}}$$

Equation 1. Boltzman Equation

P_i is the probability that cognitive agent i follows agent j according to the function M, which can be either the majority or the most influential agent. t is the temperature which determines the randomness of the process and it is set at 0.35 for convenience.

```
(p current-not-repeated-increasing-then-keep
    =visual-location>
      isa visual-location
      !eval!   (equal =current-repeated false)
==>
    +vocal>
      isa speak
      string "keep")
```

a) If current color doesn't conflict with any color of local neighborhood and persists over time then keep the same color

```
(p cur-repeated-prev-not-repeated-then-change-prev
    =visual-location>
      isa visual-location
      !eval!   (equal =current-repeated true)
      !eval!   (equal =previous-repeated false)
==>
    +vocal>
      isa speak
      string "change-previous")
```

b) If current color is repeated but the previous selection is not conflicting then change again to the previous color

```
(p cur-repeated-prev-repeated-then-change-another
    =visual-location>
      isa visual-location
      !eval!   (equal =current-repeated true)
      !eval!   (equal =previous-repeated true)
==>
    +vocal>
      isa speak
      string "change-another")
```

c) If both current and previous colors are conflicting then stochastically change to another color

```
(p finished-then-keep
    =visual-location>
      isa visual-location
      !eval!   (equal =finished true)
==>
    +vocal>
      isa speak
      string "keep")
```

d) If all cliques are in a non-conflicting state the keep color

Figure 5. Coloring: Fragment of productions to avoid local color similarities.

4.2 Activation Process

All the knowledge required for decision making is encoded in the form of ACT-R productions and the activation of those productions combines two main sub-process: Spreading activation and Partial Matching. Spreading Activation is a sub-symbolic process on which chunks spread activation to the chunks in declarative memory based on the contents of their slots. They spread an amount of activation based on their relation to the other chunks. An ACT-R feature called Production Partial Matching (PPM) is used to select the production that best matches the current sensed state of the neighborhood. Normally, a production is said to match only if the constraints specified in the *if* (condition) part match exactly to the contents of the specified buffers. With PPM enabled, the architecture will match productions in the absence of an exact match between the specification and the buffer contents. It does this by calculating a similarity value between the specification and the content. This similarity value is combined with the production's existing utility value to generate a new utility that reflects both the overall goodness of the production and the degree to which it matches the situation. This process is repeated for other productions and the production with the highest utility (after noise) is selected. In PPM, noise plays an important role in the selection of productions by simulating the subtle changes of action that humans perform according to environmental conditions and biased believes of the reality. PPM allows ACT-R to generalize the sensory input to new situations. Because the degree of match is combined with a production utility, which is itself learned from rewards reflecting its effectiveness, it provides the possibility of adaptively learning which decompositions are most effective and thus how broadly they can be generalized.

4.3 Reinforcement Learning

The reinforcement model supports the utility learning mechanism of ACT-R. The utilities of productions can be learned as the model runs based on rewards that are received from the environment. The utility of every production is updated according to a simple integrator model. If $U_i(n-1)$ is the utility of a production *i* after its n-1st application and $R_i(n)$ is the reward the production receives for its nth application, then its utility $U_i(n)$ after its nth application will be as in equation 2 (typically, the learning rate α is set at 0.2).

$$U_i(n) = U_i(n-1) + \alpha[R_i(n) - U_i(n-1)] \quad \text{Eq. 2}$$

In our experiments, cognitive agents were requested to maximize their expected total reward over a given number of trials and learn about the structure of the environment by taking into account the reward associated with each choice. Due to the fact that we run two kind of experiments, we proposed a different reinforcement algorithm for each one of these.

```
Consensus Reinforcement:
def: reward R
    if  (current_color  =  majority_color)
    then R <- R + (payoff / 100)
    #influence of current cognitive agent
    over its neighbors
    if (time_step > 0) then
        loop over neighbors
```

```
if (neighbor_color = current_color) then
R <- R + 5
if (cur_consensus > prev_consensus) then
R <- R + (payoff / 60)
else R <- R - (payoff / 20)

Coloring Reinforcement:
def: reward R
if (current_color is not in conflict)
then R <- R + (payoff / 20)
else R <- R - (payoff / 20)
if (current_color is steady and durable
decision) then R <- R + 1
    if(game_progress is increasing) then
    R <- R + (payoff / 100)
    else R <- R - (payoff / 50)
```

5. Experimentation

The purpose of the experiments is manifold: firstly it aims to compare the convergence speed and performance of both the consensus and the coloring experiments, secondly it expects to get close similar results between the cognitive simulation and the real experiments with humans and finally it seeks to find meaningful variances regarding to the different modeling strategies for decision making process (deterministic, stochastic and variable stochastic)

5.1 Convergence Speed and Performance

In figure 6 are plotted the convergence graphs for every variation of probability q for the simulated consensus experiments. Speed of convergence is increased as far as the probability q of rewiring approaches to 1.0. In figure 7 you can see that when q=0.0 the curve converges slowly and there are more fluctuations reflecting conflicts to come to an agreement whereas when q=1.0 the curve converges promptly and gently. The higher the probability q is the quicker the curve converges because of more intra-communication among the cliques is carried out and less clusters formation is observed.

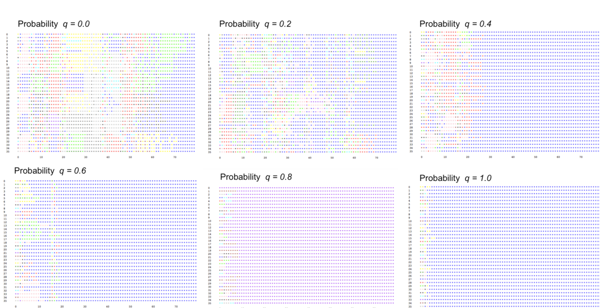

Figure 6. Color decision over time for the consensus experiments when varying the probability of rewiring, q. y-axis corresponds to time and x-axis is each one of the agents

The opposite effect occurs in the coloring experiments. Figure 8 reflects that the lower the rewiring probability is the faster and gentler the curve converges (in this case, decreasing the number of conflicts from 190 to 0 over time). It is important to notice that this opposing effect is due to the fact that when probability q is low there are not so many inter-connections between the cliques, which allows to come faster to a non-conflicting coloring situation in comparison with high clique inter-connections which require more effort to avoid repeating colors into the local neighborhood.

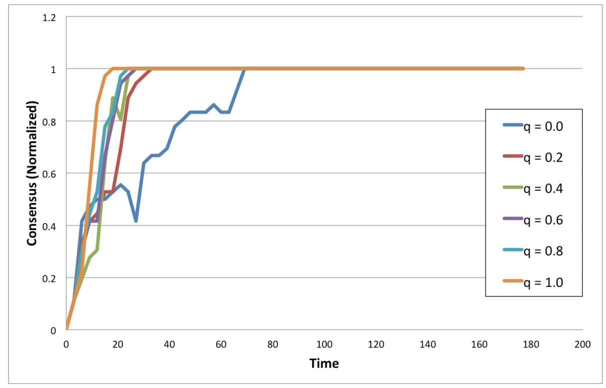

Figure 7. Convergence curve for the different variations of probability q in the consensus experiments.

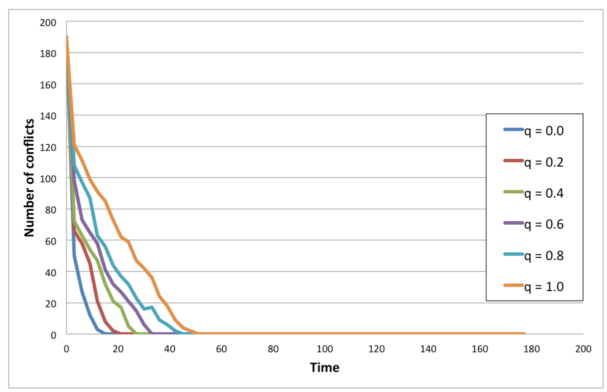

Figure 8. Convergence curve for the different variations of probability q in the coloring experiments

5.2 Consensus vs. Coloring Experiments

In figure 9 are plotted both agent-based experiments (solid lines) and human-based experiments (dashed lines). Solid lines show that our cognitive models do indeed broadly approximate the human collective behavior reported in (Kearns, 2010). In general terms, both consensus curves (agent-based and human-based) has a tendency to reduce the number of running steps to converge into a consensus when probability q increases, whereas both coloring curves tend to increase the number of running steps when probability q is increased as well. However, in our simulated experiments the cognitive mechanisms behind the decision-making accentuate the difference between running times when q = 1.0 for the coloring experiment, which emphasize the fact that a highly inter-connected cliques require more time to reach an agreement, specially if the chromatic number is low (4 in this case).

During the agent-based simulation, the gradual reduction of the chromatic number while probability q increases in coloring experiments was a crucial aspect that allowed obtaining more similar results in comparison with human-based experiments. We found that both human-based and agent-based variability was not significant in both consensus and coloring experiments when $P < 0.05$, whereas it was significant when $P < 0.001$. This variability in both experiments was improved in a second experimental phase.

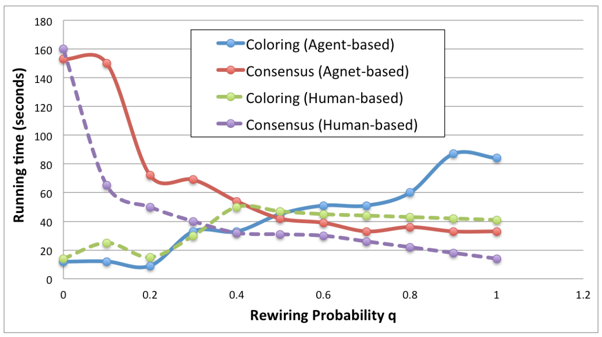

Figure 9. Coloring vs. consensus performance

It is important to notice that agent-based experiments behaves worst than human-based experiments after q = 0.5. The main reason for that is stubborn productions produces a snowball effect when the network is highly interconnected. One way to avoid that is increasing the punishing feedback received for those stubborn productions after triggering.

5.3 Multi-strategy Selection Process

We tested 3 different strategies for choosing either the most influential agent (consensus) or the next choice of a non-conflicting color (coloring). The first one was a deterministic strategy; the second one was stochastic but using a fixed value for the temperature in Boltzman equation (t = 0.35) which has demonstrated to be emerging as a reasonable setting for this parameter in previous works of ACT-R (Anderson, 2004); and the third strategy consisted of executing a set of variations for cooling the temperature slowly over time. Figure 10 summarizes the results for the consensus problem. In order to determine whether the Boltzman approaches curves were significantly different, we performed 100 runs of the experiments and collect the corresponding data.

Graphically we can conclude that the deterministic strategy keeps almost steady with a soft tendency to decrease when q increases, whereas stochastic strategies shows more fluctuations for low values of q but rapidly converging when q gradually increases. Empirically, we have found that cooling the temperature slowly from (t = 1) to (t = 0.35) in 17 cycles (time steps) was the best configuration that reflected in a more accurate fashion the human-based experiments. Furthermore, this configuration had the lowest data dispersion (with a std. dev. of 15.5). After executing an Anova test we found that $F > F_{crit}$ (2.03 > 0.77) reflecting that there is a significant difference between the strategies and, in this case, that using a variable stochastic strategy meaningfully improves the accuracy of the simulated experiments in comparison with the social behavioral experiments.

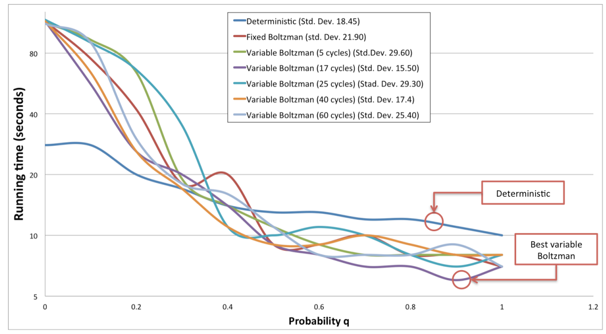

Figure 10. Performance comparison of different strategies

6. Conclusions

Opposing tensions generated by conflicting incentives over the whole cognitive process has demonstrated to properly drive the decision-making process of both individual and social levels of the multi-agent system.

Opposing productions had to compete for the right to be selected, pretty much the same that occurs in our brains when we have to make decisions, negotiate with others or come to an agreement when conflicting interests are present. In our cognitive simulation, there was supporting evidence that using stochastic strategies driven by slow decreasing of randomness fit better with the results of social experiments with humans.

Presumably it agrees with the fact that humans make less random decisions when we have more information about the dynamics of the environment, as happened with the cognitive simulation. As a final remark, modeling cognitive social behaviors is an complex task which should have into account some other aspects of human decision-making such as mood states, intentions, expectations, believes, etc. Modeling these aspects probably would probably improve the accuracy of our experiments.

Acknowledgments

This work was conducted under the Collaborative Technology Alliance Program, Agreement W911NF-10-2-0016.

3. References

Anderson, J. R., Lebiere, C (1998): The Atomic Components of Thought. Lawrence Erlbaum Associates, NJ

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., Qin, Y (2004): An integrated theory of the mind. Psychological Review 111 (4), pp 1036-1060

Colin, C. (2003): Behavioral game theory: experiments in strategic interaction.Princeton University Pres

Kearns, M., Judd, S., Vorobeychik, Y. (2010): Behavioral dynamics and influence in networked coloring and consensus. In: Nat. Ac Sc, (107), pp 14978-14982.

Kearns, M., Judd, S., Tan, J., Wortman, J (2009).: Behavioral experiments on biased voting in networks. In: Nat Ac Sc, (106), pp 1347-1352

Kearns, M., Suri, S., Montfort, N. (2006): An experimental study of the coloring problem on human subject networks. In: Science, Vol. 313, pp 8274-8278

Kirkpatrick, and C. Gelatt and M. Vecch. (1983): Optimization by Simulated Annealing. In: Science, Vol. 220, pp 671-680

McCubbins, M., Paturi, R. (2009): Networked coordination:effect of network structure on human subjects' ability to solve coordination problem. pp 899-920

Reinhard, S., Rolfi, S. (1986): End behavior in sequences of finite prisoner's dilemma supergames: a learning theory approach. In: Econ Behav Orga, pp 47-708

Russell, C., Douglas, D., Robertsi, F.. (1996): Cooperation without reputation:experimental evidence from the prisoner's dilemma games. In: Game Econ, pp 187-218

Yoella, R. Alvi.:The speed of learning in noisy games (2006): partial reinforcement and the sustainability of cooperation. In: Am Econ Rev, Vol. 96, pp 1029-1042

Author Biographies

OSCAR ROMERO is a post-doc researcher at Carnegie Mellon University. He has published mostly on Machine Learning applications and Cognitive Architectures for Intelligent Agents. His research topics focus on modeling Cognitive Phenomena and Social Behaviors.

CHRISTIAN LEBIERE is a Research Faculty directing the FMS Cognitive Modeling Group in the Psychology Department at Carnegie Mellon University. He has worked on the development of the ACT-R hybrid cognitive architecture and was co-author with John R. Anderson of the 1998 book The Atomic Components of Thought. The ACT-R cognitive architecture is used by a large international community of researchers and has been featured in hundreds of publications in the fields of Cognitive Science and Artificial Intelligence.