2014 International Joint Conference on Neural Networks (IJCNN)
July 6-11, 2014, Beijing, China

On the relationships between social structures and acquired knowledge in societies

Toshihiko Matsuka, Hidehito Honda

***Abstract*— Many existing studies on human learning pay almost exclusive attention to how individuals learn. Unlike those studies, we examined influence of social structures on knowledge acquired by societies using computer simulations. We compared four types of social networks, namely regular, random, small world, and scale-free networks. When individual differences and the principle of homophily (i.e., people who have similar beliefs tend to have close relationships with each other) exist in societies, the societies would acquire pareto-optimal knowledge. We also investigated influences of highly connected individuals on knowledge acquired by societies. The results inarguably indicate that highly connected individuals play important roles in social learning, setting the standards for what type of knowledge to be acquired by societies. *Index Terms*— ocial Learning; Social Networks; Multi-agent Simulationocial Learning; Social Networks; Multi-agent SimulationS**

I. Introduction

Many existing studies on human learning in Cognitive Science pay almost exclusive attention to how individuals learn. However, people acquire knowledge not only through individual learning, but also through interacting with others. Pentland [7] argued that influences of social structures and activities need to be considered in order to better understand true human cognitive behaviors. Likewise, Goldstone and Janssen [4] emphasized the importance of research on collective behavior. For example, they pointed out that "interacting ants create colony architectures that no single ant intends," indicating that social interactions can produce unique dynamics of knowledge acquisition that cannot be clarified by studies on individual's micro-level processes in knowledge acquisition.

In the present paper, we examine how a society as a whole acquire knowledge where each individual collaboratively learns from each other.

A. Knowledge to be learned

In the present paper, we considered coefficients of linear regression as elements of knowledge to be learned among individuals in a society, and heuristic-based optimizations of the coefficients as learning. Thus, people are assumed to learn relationships between the criterion/dependent variable and a set of predictor or independent variables. A linear regression model can be expressed as follows:

$$y_i = \sum_{j=0}^{J} b_j^{(k)} x_{ij} \tag{1}$$

here y_i is the criterion variable, i indicates a particular data point or exemplar, x_{ij} is the predictor variable j, and $b_j^{(k)}$ is person k's knowledge about relationship between the dependent variable and the jth independent variable (b_0 is being the intercept).

Although treating sets of coefficients of a linear regression model as human knowledge may seem unrealistic, what are learned in many models of human cognition are indeed numerical variables or parameters that represents "knowledge" in the models. As an initial attempt, we use one of the simplest numerical model in order to investigate dynamics of knowledge acquisition in societies. In addition, we incorporated a type of genetic algorithms in modelling social learning, thus our paradigm can be easily extended to examine models in which knowledge is symbolically represented.

II. Learning Algorithms

A. Overview of Learning Algorithm

We assumed that quite simple learning processes take place in a society. In particular, we assumed that people communicate and exchange elements of their of knowledge with others where each individual combines his or her knowledge with those of another individual. We refer to this process as "Knowledge Combination." Knowledge Combination may be interpreted as formations of new hypotheses. We also assumed that each individual has their own belief about what constitutes "good" knowledge, and knowledge that is believe to be good will be kept by individuals and therefore by the society. We refer to this process as "Knowledge Selection."

In modeling the abovementioned learning strategies, we incorporated a type of Evolution Strategy (ES) techniques in the present research. Knowledge Combination is achieved by what is called *crossover* in evolutionary computation literature in which randomly selected two individuals exchange elements of their knowledge (i.e., coefficients). There may be miscommunication in Knowledge Combination, and thus random perturbations are involved during this process. In random perturbations, which may be interpreted as *mutation*, a small random value drawn from the Normal distribution is added to each element of knowledge. After new knowledge is formed through Knowledge Combination, each individual

Toshihiko Matsuka is with Department of Cognitive and Information Science, Chiba University, Chiba, Japan (email: matsukat@muscat.L.chiba-u.ac.jp). Hidehito Honda is with Department of Cognitive and Information Science, Chiba University, Chiba, Japan.

This work was in part supported by the Japan Society for the Promotion of Science KAKENHI (Grant No. 20700235), and Nakajima Foundation

assesses his or her own knowledge on the basis of self-defined knowledge utility. Knowledge with high utility values will be kept by individuals and the society, while that with low utility values will be discarded.

B. Social structures

Four types of social structures are examined in the present research, namely, random, regular, small-world, and scale-free networks. A random network is a network where each node (person) is randomly connected with a given number of other nodes. A random network is characterized as having a short average path length (average of shortest distances among all nodes) and a low cluster coefficient (proportion of nodes, that are connected to node x, are connected to each other). In terms of a social network, a low cluster coefficients indicates that it is less likely that person x's friends are themselves friends. A regular (lattice) network is a network where each node is connected with neighboring n nodes. A regular network has a long average path length and a high cluster coefficient. Although, neither random nor regular networks are realistic model of a human society, we incorporated those networks as benchmarks.

A small world network, usually created by randomly rewiring about 10% of connections of a regular network, has a short average path length and a high cluster coefficient [8]. Previous studies have shown that many real world networks have analogous network structure to a small world network. For example, collaboration networks of film actors [8], networks of scientific collaboration [6], and ownership links among German firms [4] are shown to be structured as small world networks.

The last social network structure we examine is a scale-free network [1]. Scale-free network model incorporates both growth (number of nodes increases) and preferential attachment (probabilities that a newly added node will be connected to other nodes are proportional to the number of connections that those nodes have). Because of preferential attachment, the number of connections for each node differs greatly in scale-free networks, unlike small world networks. In a scale-free network, the numbers of connections that each node has follows a power-law distribution regardless of the size of a network. There are many nodes with smaller numbers of connections while there are only few nodes with many connections. The nodes with many connections are called "hubs." Although, existence of a hub is well known phenomenon in real world networks [1][6], among four network structures examined in the present research, hubs exist only in scale-free networks. Scale free networks also have short average path lengths and high cluster coefficients. However, cluster coefficients depend on the number of connections and its distribution also follows a power-law distribution [1].

1) Communication in a society: We assumed that people have interactions with a limited number of individuals. Only connected individuals are able communicate with each other. For regular and small world networks, we further assumed that the principle of homophily exists in a society such that people who have similar beliefs (about what constitutes "good" knowledge) would have close relationships with each other and that those who have close relationships would learn from each other. This assumption has reasonable face validity as, for example, right-wing conservatives often omit what is being stated by left-wing liberals or vise versa. In those networks people exchange information with people from their close friends, meaning that there are several more-or-less independent clusters in a society. People within the same cluster have the similar beliefs about what constitutes good knowledge, while different clusters of individuals possess different beliefs.

The principle of homophily does not exist in a random network, because its connections are random.

C. Knowledge Combinations

In Knowledge Combination, randomly selected pairs of individuals who are connected with each other exchange information to form new knowledge. The model utilizes discrete recombination for knowledge parameters . Thus,

$$b_l^{(k)} = \begin{cases} b_j^{(k)} & \text{if UNI} \leq 0.5 \\ b_j^{(m)} & \text{otherwise} \end{cases} \tag{2}$$

where UNI is a random number drawn from the Uniform distribution, and m indicates person who has a connection with person k. For self-adapting strategy parameters (i.e., σs), intermediary recombination (simple arithmetic average) is used:

$$\sigma_j^{(k)} = \frac{\sigma_j^{(k)} + \sigma_j^{(m)}}{2} \tag{3}$$

The parameters for self-adaptation (σs) are the parameters that define search widths (i.e., learning rates) for the elements of knowledge (i.e., **b**). A unique search width is allocated to each element within individuals so that sensitivity to objective hypersurface is individually tailored to meet his or her learning objectives.

This combination process continues until every individual completes forming new knowledge.

1) Inaccurate Knowledge Combination: Knowledge Combination is assumed to involve inaccurate processes, as human communication are not always perfect. Each individual's knowledge elements are randomly perturbed as follows:

$$\sigma_j^{(k)}(t+1) = \sigma_j^{(k)}(t) \cdot \exp(N(0, \gamma)) \tag{4}$$

$$b_j^{(k)}(t+1) = b_j^{(k)}(t) + N(0, \sigma_j^{(k)}(t+1)) \tag{5}$$

where t indicates time, γ defines global search width (via σ's), and $N(0, \sigma)$ is a random number drawn from the Normal distribution with the corresponding parameters.

D. Knowledge Selection

We assumed that there are two "universally" important elements in determining utility of knowledge about relationships between the criterion variable and a set of predictor variables. One is accuracy and the other is simplicity. Everyone, regardless of his or her belief about what constitutes good knowledge, evaluates his or her knowledge on the basis of those two elements. However, individuals from different clusters differently weight the importance of those two elements. In the present research we operationally define different beliefs by different sets of weight vectors

1) Knowlege Inaccuracy and Complextiy: In the model, inaccuracy (thus accuracy) of a particular set of coefficients (knowledge) is estimated based on a set of all unique data point in a training set. Thus, knowledge inaccuracy for person k is given as follows:

$$E(\mathbf{b}^{(k)}) = \sum_{i=1}^{I} \left(y_i^{(k)} - \sum_{j=1}^{J} b_j^{(k)} x_{ij} \right)^2 \tag{6}$$

where *I* is the number of unique training data exemplars, $y_i^{(k)}$ is the true value of the criterion variable for exemplar i, and the second term in the right-hand side of the equation is the predicted value for the criterion variable for person k. The desired output values are assumed to be obtained individually and thus Knowledge Inaccuracy is individually estimated.

Complexity (simplicity) of a particular set of coefficients is given as follows:

$$C(\mathbf{b}^{(k)}) = \sum_{j} \left(b_j^{(k)} \right)^2 \tag{7}$$

This complexity measure simply signifies absolute magnitudes of associations between the criterion and predictor variables. Thus, when predictor variables and the criterion variables are weakly associated, this measure tends to be small. Knowledge complexity is also estimated individually.

2) Individual Differences in Learning Objectives: Although we assumed that all individuals take both accuracy and simplicity into account in learning, there are some individual differences in weighting those two properties. We consider the differences in weights corresponds to difference in their beliefs. We define v_E as a scalar weighting for relative importance of Knowledge Inaccuracy, and $v_C = 1 - v_E$ for Knowledge Complexity.

Using these weights and Knowledge Inaccuracy and Complexity measures, we let

$$F\left(b^{(k)}\right) = v_E \frac{E\left(\mathbf{b}^{(k)}\right) - min_E}{max_E - min_E} + v_C \frac{C\left(\mathbf{b}^{(k)}\right) - min_C}{max_C - min_C} \tag{8}$$

as an overall fitness value of knowledge for a given belief (a particular Inaccuracy - Complexity weighting vector). Since knowledge inaccuracy and complexity are in different scales, they are normalized with corresponding minimum and maximum values. The minimum values are values calculated based on data created in simulations. Since theoretical maximum values are infinite for both inaccuracy and complexity, we set the maximum values as $100 \times min$.

III. Simulation

In order to explore how social interactions would produce unique dynamics of knowledge acquisition, two simulation studies were conducted. In both simulation studies, individuals in a society learns relationships between the criterion and a set of predictor variables. In Simulation 1, we examined characteristics of knowledge acquired by the societies that are organized as random, regular, and small world networks. In Simulations 2, we examined scale free networks. The main reason why we conducted separate simulations for scale free networks was that we were able to create other three networks with the same model (i.e., WS model, [8]) with different parameter values, which allowed us to control the characteristics of the simulated individuals (e.g., number of connections) in Simulation 1. Scale free networks requires a separate model (BA model; [1]).

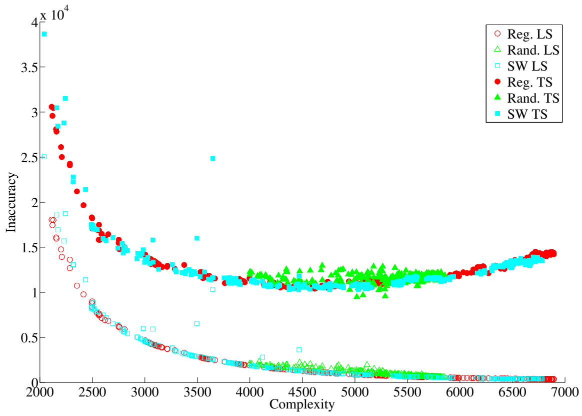

Fig. 1. The characteristics of knowledge acquired by regular, random, and small world networks for a particular data set. LS and TS indicate learning and testing data sets, respectively. In regular and small world networks some individuals acquired very accurate knowledge at the cost of complexity, while others acquired very simple knowledge at the cost of accuracy, showing that those societies as a whole formed Pareto-optimal knowledge. Knowledge acquired by a random network was less diverse than those of regular and small world networks. Most individuals in a random network acquired generalizable knowledge (i.e., knowledge inaccuracies for new data were relatively low).

A. Simulation 1

1) Method: There were a total of 50 predictor variables. Among them, only 20 variables have meaningful associations with the criterion variable. In other words, the criterion variables were created with 20 predictor variables and random noise drawn from the Normal distribution.

$$y_i = \sum_{j=1}^{20} b_j x_{ij} + N(0, 100) \tag{9}$$

A total of 100 data points were created in each simulation, among which 50 data points were used for learning and the remaining 50 points were used for testing/generalization. min_C was defined as $\sum_{j=1}^{20} b_j^2$ and min_E as $\sum_i (y_i - \sum_{j=1}^{20} b_j x_{ij})^2$

The model was run in a simulated social learning procedure with 1000 generations (communication) to learn the relationship between the criterion and predictor variables. The model parameter was arbitrary selected (γ = 0.1). There were a total of 300 individuals in a society. Each individual

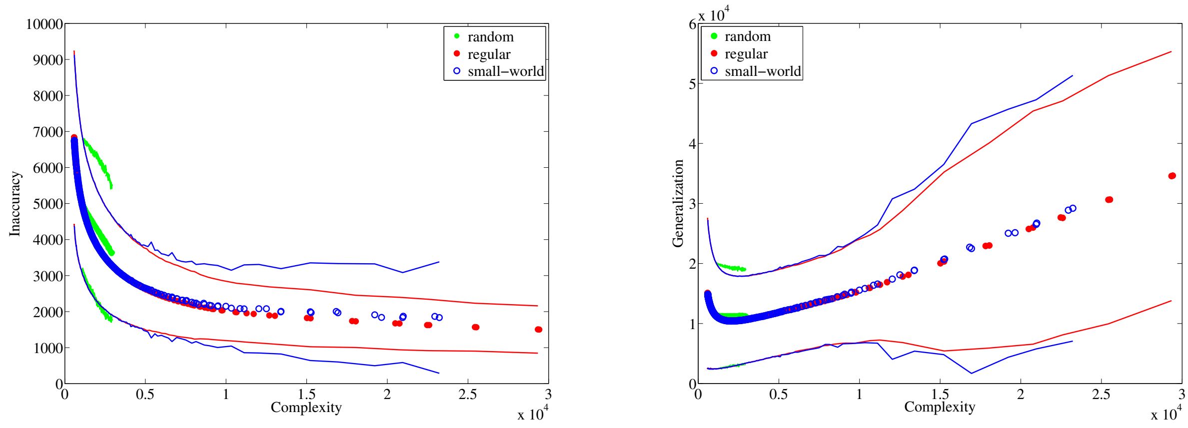

Fig. 2. Results of Simulation 1. Left panel shows knowledge complexity-vs-inaccuracy trade-offs, where each dot represents average of single individual, and lines indicating $\pm$ 2 standard deviations. Right panel shows knowledge complexity-vs-generalization trade-offs.

had exactly 10 connections. The scalar weights that define relative importance for Knowledge Accuracy (i.e., v_E) were evenly spread from 0 and 1 for 300 individuals. Note that the weight for Knowledge Complexity was 1 minus Knowledge Accuracy ($v_C = 1 - v_E$). There was a total of 300 simulations for all three social network structures.

2) Results and Discussion: Figure 1 shows the characteristics of knowledge acquired by regular, random, and small world networks for a particular data set, for both learning and testing data sets. Figure 2 shows characteristics of knowledge acquired by individuals in societies, where each dot represents knowledge acquired by one individual (averaged). The left panel shows relationship between knowledge inaccuracy and complexity, and the right panel for knowledge generalization (inaccuracies for new data sets) and complexity. The figures and Table 1 show that there was a great degree of individual differences in acquired knowledge in regular and small world networks. Some individuals acquired very accurate knowledge at the cost of complexity, while others acquired very simple knowledge at the cost of accuracy.

The figures also shows that the regular and small world-like society as a whole formed thorough Pareto-optimal knowledge. That is, it is very less likely that one individual's acquired knowledge was simultaneously better in both accuracy and simplicity than those of other individuals.The results can be interpreted as that those societies would acquire cluster of knowledge that exceed at least one important aspect of knowledge when there are individual differences in beliefs and when individuals learn from others who share similar beliefs and values. This result was not surprising, because social learning processes that take place in regular and small world networks resemble one of multi-objective evolutionary optimization methods called vector evaluated approach (Deb, 2001). The resemblance may indicate that the principle of homophily (i.e., people who have similar beliefs tend to have close relationships with each other) and individual differences together can lead a society to acquire and hold pareto-optimal knowledge.

Individual differences alone is not sufficient for describing acquisition of thorough pareto-optimal knowledge, as the random networks did not form such knowledge. Knowledge acquired by a random network society was less diverse than those of regular and small world networks. Random connections might have canceled out individual differences. However, most individual in random networks acquired generalizable knowledge (i.e., knowledge inaccuracies for new data were relatively low). In fact, the random network resulted in the lowest average generalization error among the three social networks. In contrast, some individuals in regular and small networks acquired either under or over generalizing knowledge.

B. Simulation 2

1) Method: In Simulation 2, we examined knowledge acquisition in scale-free networks. In particular we pay close attention to influences of nodes with many connections (i.e., hubs) on overall acquired knowledge in societies.

The general procedures of Simulation 2 follow those of Simulation 1. In Simulation2, the same data sets (i.e., ys and xs) were used and same knowledge acquisition processes were applied as in Simulation 1.

Because it is quite difficult to control dispersion of beliefs (about what constitutes good knowledge, i.e., v_E & v_C) with a scale-free networks, we randomly permutated equally spaced v_Es among individuals.

2) Results and Discussion: Given that v_Es were randomly assigned to individuals, inaccuracy-complexity trade-off of acquired knowledge in scale free networks resembled that of random networks

Figure 3 shows the relationships between knowledge characteristics of the node with the highest number of connections and those of other individuals within the same society. The relationships between the hub's and average knowledge inaccuracy in learning was quite strong. Their correlation coefficient was 0.744. The relationship between the hub's and average knowledge complexity was also strong, and its correlation coefficient was 0.747. The relationship between

TABLE I

DESCRIPTIVE STATISTICS OF KNOWLEDGE ACQUIRED BY SOCIETIES

Type	Inaccuracy		Complexity		Generalization	
	Avg.	Std.	Avg.	Std.	Avg.	Std.
random	4389.1	590.9	1869.4	813.6	1114.2	810.2
regular	4231.5	1530.4	3324.6	4708.4	1268.1	4008.1
small world	4245.6	1489.9	3155.9	4152.2	1252.5	3567.4

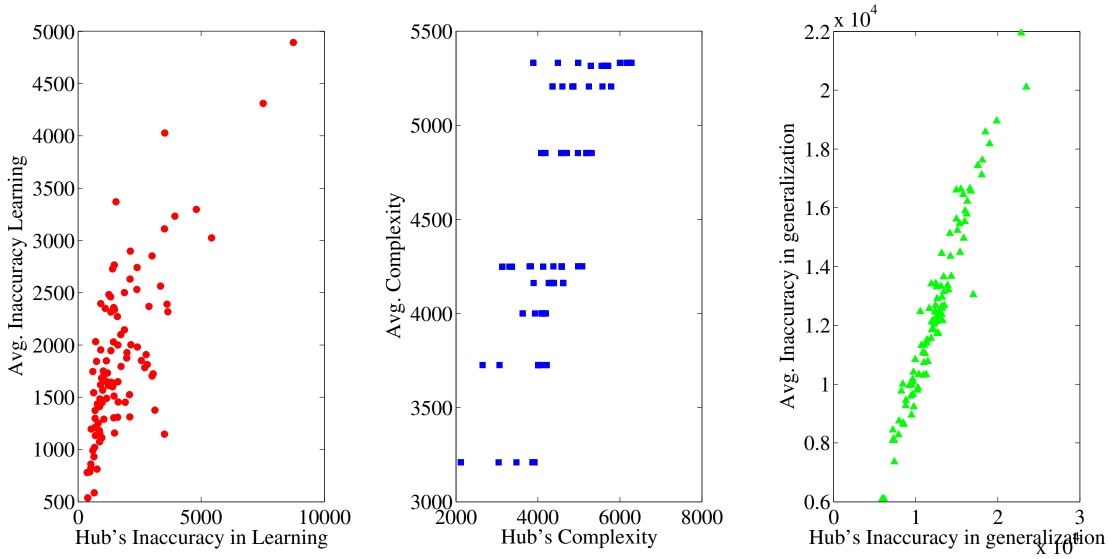

Fig. 3. Results of Simulation 2. The left panel shows relationship between knowledge inaccuracy (in learning) of the node with the highest number of connections (i.e., hub) and average knowledge inaccuracy among individuals in the same society. The middle and right panels show relationship between knowledge complexity and generalization inaccuracy of the hub and average knowledge complexity and generalization inaccuracy, respectively.

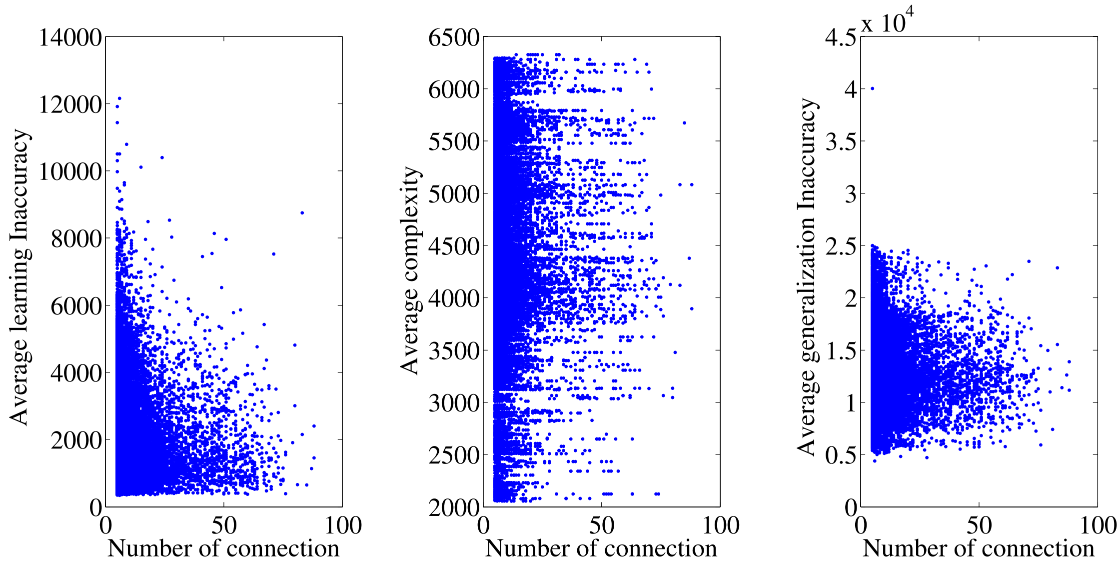

Fig. 4. The left panel shows relationship between nodes' connectivities and knowledge learning inaccuracy. The middle and right panels show relationship between nodes' connectivities and knowledge complexity and generalization inaccuracy, respectively.

hub's and average generalization inaccuracy was remarkably strong. Its correlation coefficient was 0.970.

These results inarguably indicate that the hubs play important roles in social learning. They set standards for knowledge complexity and knowledge accuracy. However it is quite difficult to understand why the hubs' knowledge generalization inaccuracies have stronger associations with that of other individuals as compared with knowledge inaccuracies (for learning sets) and complexities. Figure 3 shows the relationships between nodes' connectivities and knowledge accuracies, complexities, and generalization. Higher numbers of connections were associated with lower learning inaccuracies and intermediate levels of generalization inaccuracies. There was no clear relationship between the number of connections and knowledge complexities. It also indicates that there were greater levels of variabilities in knowledge complexities than learning inaccuracies. This in term suggest that optimization of knowledge complexity was more difficult than optimizing knowledge inaccuracy (or max_C was set too large or more learning generation might have been needed). A wider range of knowledge complexities in societies more or less evened out the societies' needs for simplifying knowledge, and then they acquired knowledge at middle levels of complexities. If this is the case, the results of Simulation 2 may not be caused by the nature of scale-free networks, but by the one specific instantiation of scale-free networks. Further simulations and analyses are needed to examine this issue.

IV. Conclusion and Future Directions

Many existing study on human learning in Cognitive Science pay almost exclusive attention to how individual learns. In the present research, unlike previous studies, we examined learning processes that take place in society using computer simulations. We examined how society as a whole acquired knowledge while each individual interacts with others. We compared four types of social networks, namely regular, random small world, and scale-free networks. Where applicable, we assumed that the principle of homophily (i.e., people who have similar beliefs tend to have close relationships with each other) and individual differences exist in societies. In such societies (individual differences and homophilic), we found that the society would acquire pareto-optimal knowledge, such that there is no cluster of knowledge that was worse (or better) in two important aspects of knowledge (i.e., accuracy and simplicity) as compared with those of other clusters. That society acquired very robust and wide variety of knowledge. In Simulation 2, we investigated influence of highly connected individuals on knowledge acquired by societies. The results inarguably indicate that highly connected individuals play important roles in social learning, setting the standards for knowledge complexity and accuracy. Though it is still inconclusive, knowledge of highly connected individuals seem to have stronger influence on generalizabilities of acquired knowledge (vs. knowlege complexities and learning inaccuracies).

With two simulation studies, we showed that social interactions can produce unique dynamics of knowledge acquisition that is difficult if not possible to be clarified by studies on individual's micro-level processes in knowledge acquisition.

There are several ways to extend our research paradigm. In the present research, we assumed that people exchange elements of knowledge that are being heuristically optimized by individuals (i.e., **b**). In reality, people also communicate their inferences about the criterion (i.e., y in our simulations), and then optimize their knowledge on the basis of other people's inferences. Another natural and appealing extension is to incorporate unsupervised learning to see how knowledge is self-organized.

References

[1] Barabsi, A. & Albert, R. (1999). Emergence of scaling in random networks. *Science, 286.* 509512.

[2] Deb, K. (2001). *Multi-objective optimization using evolutionary algorithms*, Chichester: Wiley.

[3] Feldman, J. (2003). The simplicity principle in human concept learning. *Current Directions in Psychological Science. 12* 227-232.

[4] Goldstone, R. L., & Janssen, M. A. (2005). Computational models of collective behavior. *Trends in Cognitive Sciences, 9*, 424-430.

[5] Kogut, B., & Gordon, W. (2001). The small world of Germany and the durability of national networks. *American Sociological Review*, 66, 317-335.

[6] Newman, M. E. J. (2001). The structure of scientific collaboration networks. *Proceedings of the National Academy of Sciences*, 98, 404-409.

[7] Pentland, A. (2007). On the Collective Nature of Human Intelligence. *Adaptive Behavior, 15*, 189-198.

[8] Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature, 393*, 440-442.