Collective Cognition: Exploring the Dynamics of Belief Propagation and Collective Problem Solving in Multi-Agent Systems

Paul R. Smart[1], Trung Dong Huynh[1], Dave Braines[2], Katia Sycara[3] and Nigel R. Shadbolt[1]

[1]School of Electronics and Computer Science, University of Southampton, Southampton, SO17 1BJ, UK.
[2]Emerging Technology Services, IBM United Kingdom Ltd, Hursley Park, Winchester, SO21 2JN, UK
[3]Robotics Institute, School of Computer Science, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213-3890, US.

Introduction

We tend to think of cognition as something that takes place inside the heads of individual human agents. However, it is clear that many of our cognitive accomplishments depend just as much on our ability to exploit the elements of our social and technological environments as they do on the information processing dynamics of the biological brain (Smart et al., 2010). Our social networks constitute a particularly potent source of bio-external scaffolding: one that shapes, constrains and influences the profile of much of our daily cognitive activity. However, the precise way in which networks enable a group of agents to coordinate their thoughts and actions in cognitively-productive ways is still something of which we have, as yet, very little understanding. This paper is an attempt to review the status of our current understanding of network-enabled collective cognition and to explore ways in which our current understanding might be improved. The primary targets for discussion are the dynamics of belief propagation and collective problem-solving in multi-agent systems. These phenomena, we suggest, provide potent examples of collective cognition in terms of both cognitive state fluctuations (belief propagation) and cognitive processing (collective problem solving). In addition to reviewing the literature in these areas, we also present a number of ideas to help guide future research efforts.

Belief Propagation in Heterogeneous Agent Communities

Imagine a situation where a physically distributed group of individuals are attempting to acquire some common understanding of an event, say a humanitarian disaster. Some individuals may have direct access to information about the event (either through personal experience or via remote sensing capabilities). Others (perhaps the majority) will have to rely on the information they receive from their colleagues and co-workers (i.e. the elements of their social network). In both cases, the information received by a particular individual is likely to vary in terms of its accuracy, scope and quality, and conflicting information is likely to be an all too common occurrence. How will agents' beliefs about the focal event evolve in this situation? Will all the agents eventually converge on a common interpretation or understanding of the event, or will distinct interpretations be the norm? What, moreover, of the role played by the social/communication network in shaping the overall profile of agents' beliefs about the event? Do some network structures lead to greater convergence in agents' beliefs, and, if so, can this inform our understanding of how some agent communities arrive at a common understanding or interpretation of some external state of affairs?

In order to begin making progress on these issues, Glinton et al (2010) report the results of a number of studies that seek to examine the dynamics of belief propagation in networked multi-agent systems. Their studies rely on computer simulations involving a large number of agents – with most agents receiving information about some external state of affairs via their connected peers and the remainder receiving information from a combination of other agents and the external environment (see Figure 1). Information from the environment is provided to the agent network via a sensor, which simply returns a value of true or false reflecting the state of some fact (*F*) about the world. The task of the agents is to establish beliefs about *F* by processing information they receive either from the sensors or from other agents. Importantly, the information received from sensors is not always accurate: noise is introduced into the sensor readings such that the sensor sometimes returns an incorrect value based on a predetermined probability.

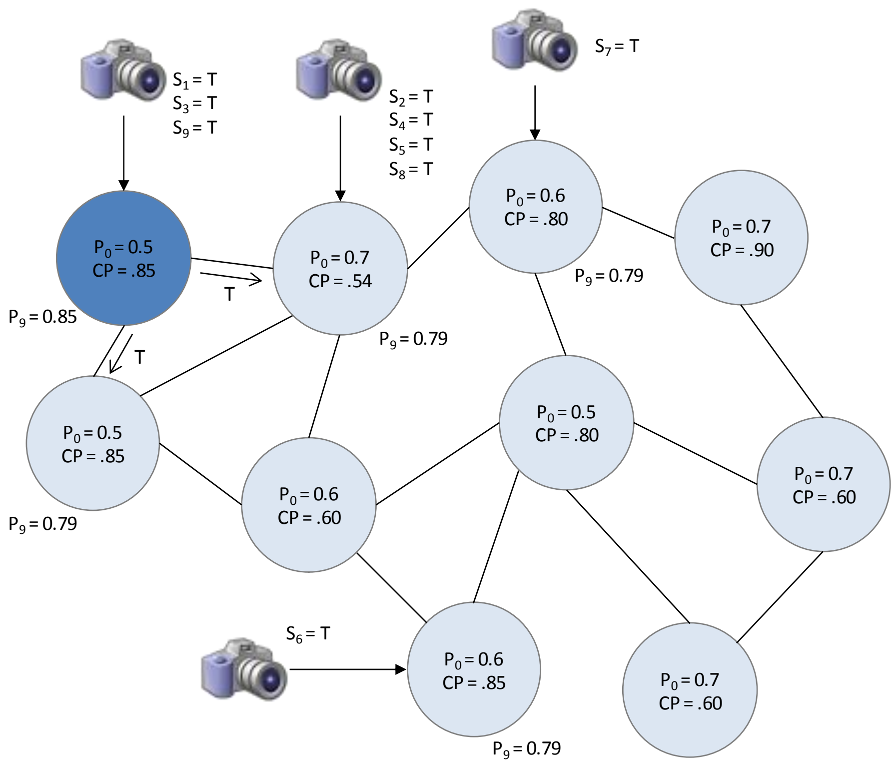

Figure 1: Simplified agent network illustrating the dynamics of belief propagation. Some of the agents are directly connected to sensor devices (represented by the camera icon), while others receive information solely from network neighbours. The figure shows the system in a critical state. Sensor readings are obtained at nine successive points in time (s_1-s_9), and this is sufficient to cause one agent (the darkly shaded node) to adopt a belief state of 'true'. The arrows indicate that the belief state value is communicated to neighbouring agents, but these neighbouring agents do not change their beliefs. The critical threshold for belief change in this system is 0.80. Thus the value of P needs to exceed 0.80 before an agent will change their beliefs. Figure adapted from Glinton et al (2010).

The agents in the simulation compute their belief in *F* based on equations 1 and 5 (see below). These equations determine the agent's belief that *F* is true given a particular sensor reading (equation 1), or that *F* is true given the value of *F* communicated by other agents (equation 5).

$$P'(b_{ai} = true) = \frac{A}{B + C} \tag{1}$$

$$A = P(b_{ai} = true)P(s_{ai} = false|F = true) \tag{2}$$

$$B = \left(1.0 - P(b_{ai} = true)\right)_t P(s_{ai} = false|F = false) \tag{3}$$

$$C = P(b_{ai} = true)P(s_{ai} = false|F = true) \tag{4}$$

$$P'(b_{ai} = true) = \frac{D}{E + G} \tag{5}$$

$$D = P(b_{ai} = true)P\left(b_{aj} = false|F = true\right) \tag{6}$$

$$E = \left(1.0 - P(b_{ai} = true)\right)P\left(b_{aj} = false|F = false\right) \tag{7}$$

$$G = P(b_{ai} = true)P\left(b_{aj} = false|F = true\right) \tag{8}$$

In these equations, $P(b_{ai})$ gives the prior belief of agent *ai* in the fact *F*, and $P(s_{ai}|F)$ gives the probability that the sensor will return a given estimate of the fact (true or false) given the actual value of *F*. $P(b_{aj}|F)$ is referred to as the conditional probability (CP). It provides a measure of the credibility that an agent *ai* assigns to the value of *F* received from a neighbouring agent (*aj*).

Processing proceeds by making the sensor readings available to the agent network and updating the agents' belief states accordingly. A threshold is used to determine whether an agent believes (at any given point in time) whether *F* is true, false or unknown. And whenever agents change their belief from one state to another (e.g. from unknown to true), the state change is communicated to all the agent's neighbours in the network.

What Glinton et al (2010) observed in these simulations is that for a particular range of CP, the dynamics of belief propagation within the agent community exhibits the properties of a Self-Organized Critical (SOC) system. That is, the number of agents who converge on a common belief state at any given point in time (switching their beliefs to assume a common state of belief in *F*) is characterized by a power law probability distribution, with many small, localized cascades of belief change punctuated by much more infrequent system-wide changes. What seems critical to the emergence of this behaviour is the value of CP, or the credibility that agents assign to the information they receive from other agents. When CP is too low, agents will almost never be influenced by the decisions of their neighbours, and thus no propagation of belief states occurs. When CP is too high, the extent of social influence is so high that any given agent may be 'persuaded' to change its belief based on the input it receives from a single neighbour. In this situation, large chains of belief propagation do occur, but such events are so frequent they violate the statistical characteristics of a SOC system.

As Glinton et al (2010) recognize, the dynamics of belief propagation in the system are determined by a combination of the decision threshold and the specific value of CP. These factors combine to provide a local resistance to belief change. Pretty much as the friction between sand grains contributes to the SOC dynamics of avalanches in an ever-growing sandpile (Bak et al., 1983), so too the resistance to social influence (i.e. the resistance to belief state changes following communication with connected peers) seems to determine the emergent profile of cognitive state changes in a community of interacting agents:

> *"The large avalanches occur because...P(baj|F) acts as a local resistance to changing belief when receiving a neighbours belief...In certain circumstances given the data received by agents thus far, many agents will simultaneously reach a state where they are near the threshold. In this case a single incoming piece of data to any of the agents will cause an avalanche of belief changes."* (Glinton et al., 2010)

Figure 1 and Figure 2 exemplify this phenomenon. In the figures, P_0 gives the prior (initial) belief of each agent and S_t corresponds to a sensor reading received at time t. Figure 1 shows a state where many agents are near the critical threshold for changing their beliefs as a result of a series of sensor readings S1...S9. Figure 2 shows the effect of adding one additional sensor reading, S10. There is a cascade of cognitive state changes as each agent changes their belief about *F*. The arrows in Figure 2 reflect the sequence of changes as each agent in the network switches its belief to coincide with the belief states adopted by all other agents.

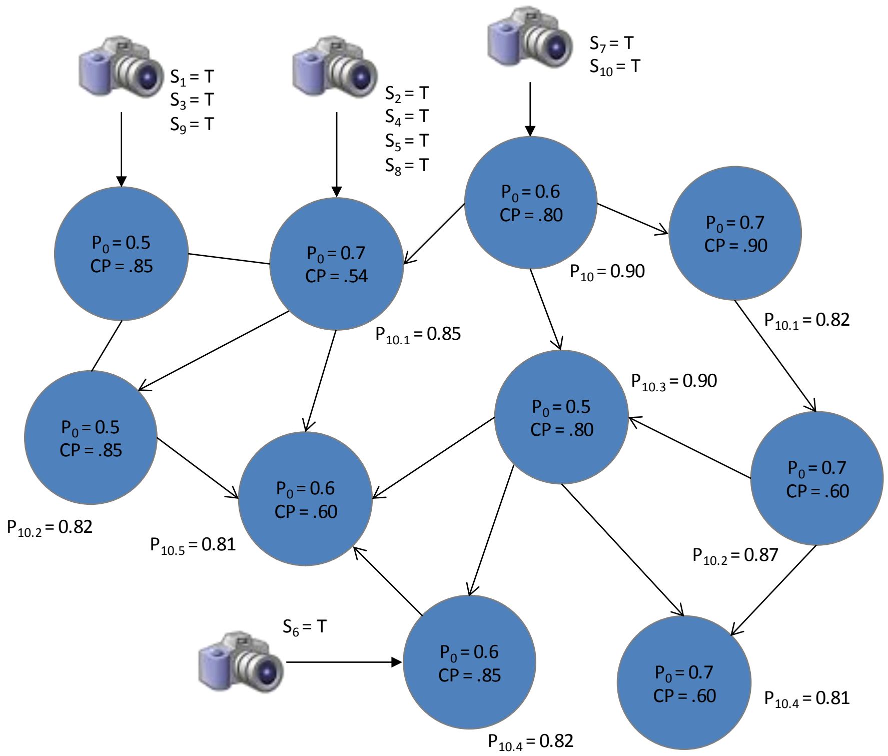

Figure 2: The same network as in Figure 1, this time following the addition of a single sensor reading (s_{10}). The sensor reading causes a cascade of belief state changes, with all agents changing their belief to 'true'. The arrows in this figure indicate the flow of information around the network. Figure adapted from Glinton et al (2010).

We can now begin to see why the cognitive dynamics of a networked multi-agent system might be profoundly affected by exposure to limited subsets of information. Providing we approach the kind of conditions modelled by Glinton et al (2010), the scale of cognitive state changes need not be in proportion to the apparent 'significance' of information to which (at least some) members of the agent community are exposed. Instead, large-scale (community-wide) changes in cognitive state may occur in response to the receipt of seemingly trivial pieces of additional information (for example, a single sensor reading in the case of Glinton et al (2010)). Inasmuch as the findings of Glinton et al (2010) reflect the processes of belief change in social networks, such studies highlight how factors like resistance to social influence and the degree of inter-agent trust (corresponding to CP in Glinton et al (2010)) might impact the 'effective' cognitive significance of external information, particularly in terms of its ability to precipitate large-scale, discontinuous transitions in the cognitive states of multiple (perhaps all) agents.

One factor that still needs to be considered here concerns the impact of network structure on belief dynamics. In one experiment, Glinton et al (2010) organized the agents into networks with different topologies (scale free, small world, hierarchical, and so on). They then examined the effect of these topologies on the extent of belief state changes. What they observed was that, for each type of network, as the value of CP increased, the size of belief cascades also increased. Thus, the effect of increasing levels of CP was largely the same across all network types, although the effect was particularly pronounced for the scale-free and small world networks and attenuated for the hierarchical networks.

In an important extension of Glinton et al's (2010) work, Parachuri et al (2010) attempted to assess the effect of including a model of human agents in the network simulations. A human agent was distinguished from a synthetic agent by assigning them a much larger CP value and reducing the frequency of belief updates. Human agents thus updated their beliefs less frequently, and when they did update (and change) their belief, their influence was significantly greater than that exerted by a conventional synthetic agent. The basis for these differences is the assumption that (in the case of elevated CP) "humans are good at analyzing (given particular information) and hence agents believe in humans more", and (in the case of reduced frequency of belief updates) "humans are slower to sense the environment or update their beliefs".

Following the integration of human agents into the simulation, Parachuri et al (2010) report what they refer to as an 'enabler-impeder effect'. This means that the progressive addition of human agents results in an initial facilitation of belief propagation (when the proportion of human agents to synthetic agents is relatively low), followed by an inhibition of belief propagation as the proportion of human agents increases. In accounting for this effect, Parachuri et al (2010) comment that the "enabler effect is due to the fact that agents have a higher belief in humans while the impeder effect is due to the fact that humans have a high latency between successive belief calculations".

What these results highlight is that differences between agents in a social network can profoundly affect the dynamics of cognitive state changes at the collective level. Although factors such as the structural organization of the network are commonly seen as crucial in terms of our understanding of network behaviour, the findings of Parachuri et al (2010) highlight the potential significance of compositional factors. Thus, the key manipulation introduced by Parachuri et al (2010) was to exacerbate the extent of inter-individual differences within the agent community, thereby making the composition of the network more heterogeneous. This heterogeneity (manifest in the differential credibility assigned to agent communications) was shown to influence the extent to which aspects of systemic behaviour (belief cascades) could be observed. The results are thus an important reminder of the potential significance of individual differences to the emergence of collective phenomena. Although much work in network science tends to overlook differences at the nodal level, choosing instead to create homogeneous networks of rather simple nodal elements, it is clear that many real-world networks are highly heterogeneous in terms of their composition. Human social networks, for example, consist of people who may differ along a number of important social and psychological dimensions (e.g. social status, authority, susceptibility to social influence, and so on). What the results of Glinton et al (2010) and Parachuri et al (2010) demonstrate is that we would probably do well to at least consider such differences if we are ever to get a grip on the collective cognitive dynamics of real-world social networks.

Modelling the Dynamics of Collective Cognition: A Network-Based Approach to Socially-Mediated Cognitive Change

Towards a Naturalistic Model of Socially-Mediated Cognitive Change

Developing an interesting model of network-enabled collective cognition arguably requires two essential ingredients: a healthy dose of psychological realism and a rather liberal spicing of real-world simplification. Any network model clearly needs to make some simplifying assumptions about the real world; otherwise it loses the elements of computational tractability and explanatory concision that make it useful as an aid to both analysis and comprehension. A network model should also (potentially at least) avail itself of a degree of psychological realism. Psychological realism is important whenever aspects of human psychology can be seen to change the dynamics of information flow and influence in a social network. Thus we saw, in the case of the study by Parachuri et al (2010), that individual differences can sometimes exert a significant impact in terms of the profile of cognitive state changes at the systemic level, and this highlights one way in which many conventional network studies (which countenance the use of largely undifferentiated nodal elements) may fall short as models of collective cognitive flux in real-world social contexts. That said, it should be clear that the studies of Parachuri et al (2010) (and Glinton et al (2010)) also adopt very simplistic models of belief propagation, and these models are, in all likelihood, inadequate when it comes to understanding the dynamics of socially-constrained belief propagation[1]. Although the studies of Glinton et al (2010) and Parachuri et al (2010) provide some initial insight into the dynamics of belief propagation in heterogeneous agent communities, the studies use a highly simplified model of belief reflecting agents' beliefs about the veracity of a particular fact (e.g. '*F* is true'). This situation is unlike that encountered in more naturalistic[2], real-world settings. In most real-world contexts, agents will possess a variety of beliefs, and the truth or falsity of any particular belief will, in part, be determined by the 'logical' inter-dependencies between the beliefs. To illustrate this point, consider a situation where a human decision-maker is attempting to formulate a comprehensive picture of a situation based on a body of incomplete and uncertain information. Suppose that the individual has already received some information that leads her to form certain beliefs. Now, whenever new information about the situation is received, the new information will be evaluated with respect to those pre-existing beliefs. In cases where a conflict is encountered (i.e. new information contradicts a prior belief), the individual will either have to revise her pre-existing beliefs or discount the newly received information. The point here is simply that whenever an agent

[1] This is, of course, not quite a case of being 'hoisted with one's own petard'. Both Glinton et al (2010) and Parachuri et al (2010) study networks that are comprised largely of synthetic rather than human agents. As a result, issues of psychological fidelity are not their primary concern, and the tenability of their conclusions are not undermined by a failure to address such issues.

[2] Just as the paradigm of naturalistic decision making (Zsambok & Klein, 1997) emphasizes the disparity between conventional cognitive models of decision-making and the reality of much real-world cognition, so the current discussion emphasizes the mismatch between conventional network models of collective cognition and much of the reality of socially-mediated cognitive change.

entertains multiple beliefs, their background (e.g. causal) knowledge often enables them to detect inconsistencies between belief states, and such inconsistencies can influence decisions about whether new information is accepted or rejected. In cases where conflicting information is accepted, the agent will be forced to revise existing beliefs (or at least downgrade the certainty assigned to those beliefs). And in cases where such information is rejected, we can begin to understand its lack of influence in terms of an existing complex of inter-dependent belief states.

The inter-dependencies between an individual's beliefs lead to what may be called states of variable cognitive consistency. A state of high cognitive consistency is one in which a set of inter-dependent cognitions (beliefs, values or attitudes, or whatever) are highly compatible and mutually reinforcing. A state of low cognitive consistency, in contrast, is one in which we encounter a set of largely incompatible or conflicting cognitions. Suppose, for example, that an individual's belief system is represented as a collection of 5 separate beliefs, each of which is modelled as a binary variable (where 1 represents a situation where an agent believes some fact is true, and 0 represents a situation where the agent believes some fact is false). In many cases, we will encounter dependencies between the various beliefs, as determined by the logical or causal structure of the domain to which the beliefs apply. We can represent these dependencies via a set of rules that indicate how a change in one specific belief is likely to influence the other beliefs to which it is connected. Thus, in Figure 3, we can see that if an individual believes that A is true, then it is likely that B is true and C is false. If the individual received reliable information at a later date that C was, in fact, true, then accommodating that information into the individual's belief system would lead to a state of low cognitive consistency. We can see that for any system of beliefs, which is modelled in this way, we can enumerate all the system states that are consistent (i.e. states in which none of the beliefs violates the dependency rules), and we can also calculate the relative consistency of different system states based on the extent to which the dependency rules are actually violated. As we will see below, this manner of modelling belief systems and representing the linkages between various beliefs is likely to be relevant to the development of computational models dealing with belief system dynamics at the societal or cultural level.

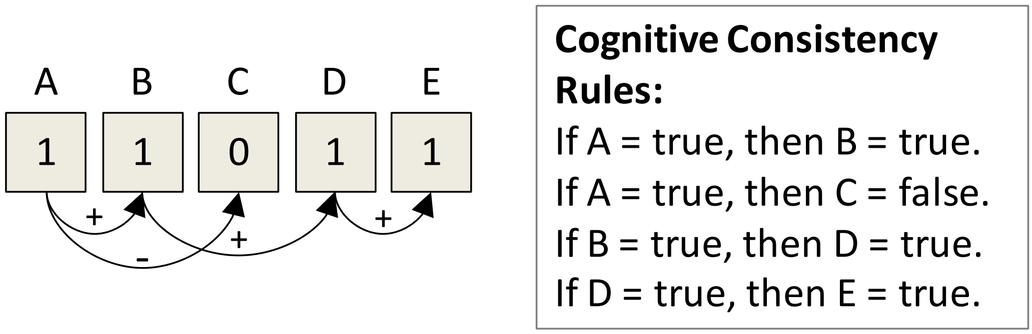

Figure 3: A primitive belief system exhibiting logical dependencies between specific belief states. For any given belief system, we can enumerate all the configurations (combinations of belief states) that satisfy the consistency rules and are thus cognitively consistent. We can also compute the relative consistency of different configurations based on dependency violations. Note that one extension to this model is to assume that the dependencies are probabilistic rather than all-or-none. For example, we could associate a specific probability with the rule 'If A = true, then B = true' to express the level of certainty or confidence in the dependency expressed by the rule.

Cognitive Dissonance

All this talk of cognitive consistency should strike a chord with those familiar with Festinger's (1957) theory of cognitive dissonance. Cognitive dissonance theory is a highly popular and influential theory in social psychology, which maintains that individuals are under internal pressure to avoid inconsistent or contradictory cognitions. According to dissonance theory, whenever an individual perceives an inconsistency in their cognitions, they encounter negative mental states (e.g. anxiety, guilt, shame, etc.), and these motivate the individual to attempt to reduce the inconsistency. When people's cognitions are consistent with one another (i.e. high states of cognitive consistency), then they are in a state of harmony or cognitive consonance.

Now, there are clearly some important parallels between the foregoing discussion of cognitive consistency and the social psychological notion of cognitive dissonance. Based on the dependencies between a variety of individual belief states, we can calculate the extent to which a particular belief system is in a state of high cognitive consistency (consonance) or low cognitive consistency (dissonance). We can additionally assume that agents will come under internal pressure to avoid dissonant cognitions – they will attempt to avoid states of cognitive dissonance, either by adjusting their specific profile of beliefs (engaging in cognitive change), or by rejecting information that would lead to dissonant cognitions (discounting disconfirming evidence)[3]. Given the widespread influence and application of dissonance theory to a variety of areas of psychology, any model of socially-mediated cognitive change is likely to benefit (in terms of psychological realism) by embracing notions of cognitive consistency. This, then, is one way in which the psychological fidelity and plausibility of computational models of socially-mediated cognitive change can be improved. Previous attempts to model the cognitive dynamics of social groups tended to adopt models that made a number of simplifying assumptions about the complexity of human cognition. Such assumptions were clearly necessary in order to make initial inroads into what is a highly difficult and complex problem area. However, if we are to further our understanding of how the collective cognitive profile of a group of human agents emerges in response to the vagaries of information flow and influence in a variety of networked situations, then it is important we factor in some of the real-world features that contribute to the stability (or instability) of cognitive states at the individual level. The notion of cognitive consistency, with its obvious linkages to cognitive dissonance, provides, we suggest, the basis for more ecologically-realistic models of socially-mediated cognitive change.

Cognitive Consistency, Epistasis and Boolean Networks

How can we develop more naturalistic, psychologically-plausible models of collective cognition without reneging on the computational-tractability required for computer simulation? Based on the foregoing discussion, we suggest that the agents in such models need to be considerably more complex than those typically employed in previous network science studies (Glinton et al., 2010; Watts, 2002). In particular, we suggest that we need to represent the inter-dependencies and linkages between a variety of beliefs, and we need to be able to compute a measure of cognitive consistency that (at least in some cases) provides the impetus for internally-driven cognitive change. In addition to this, we need to be able to represent the influences exerted by other agents in a social

[3] Of course, the strategies pursued by humans to reduce cognitive dissonance are many and varied. They may include rationalization, discounting of conflicting information, cognitive change (changes in attitudes, beliefs and values) and the expression of ego defence mechanisms.

network. We have to accommodate the fact that an individual agent may adopt a particular profile of beliefs in response to the beliefs adopted by its neighbours. Furthermore, we have to recognise the potentially complex interactions between the (internal) forces of cognitive consistency and the (external) forces of social influence. An individual may be differentially resistant to social influence depending on their current cognitive state (states of high internal cognitive consistency may make an individual relatively invulnerable to social influence), and the cognitive dynamics of the system as a whole (the cognitive trajectories pursued by all the agents in the network) may be largely governed by the inter-play between these two (occasionally opposing) forces.

One way of making progress on the model development front is to consider the kind of approaches used in the modelling of complex adaptive systems, particularly the kind of approaches used to model processes of evolutionary and co-evolutionary change. We saw, in the previous section, that the inter-dependencies between the various beliefs of an individual agent can be represented as a network of links that exert positive (reinforcing) and negative (contradictory) influences on the various beliefs to which they are connected (see Figure 3). The resulting system, comprising a network of inter-linked binary variables, is somewhat similar to the notion of Boolean networks as described by the complexity theorist Stuart Kauffman (1993, 1995). According to Kauffman (1993) Boolean networks are:

> *"systems of binary variables, each with two possible states of activity (on and off) coupled to one another such that the activity of each element is governed by the prior activity of some elements according to a Boolean switching function." (pg. 182).*

Boolean networks, Kauffman argues, enable us to explore the role of self-organization in the evolution of complex biological systems. He identifies a number of regimes in which such networks can operate, including the ordered, chaotic and complex regimes, with self-organization emerging as a feature of the complex regime. In the complex regime, writes Kauffman (1993), we encounter a phase transition which is poised between order and chaos. In this transition region "altering the activity of single unfrozen elements [network nodes] unleashes avalanches of change with a characteristic size distribution having many small and a few large avalanches" (pg. 174)[4].

The thing that determines whether a Boolean network lies in one or other of these regimes is the number of links between the nodes. Thus Kauffman (1993, 1995) shows that as the number of linkages or dependencies between the binary elements increases, the network becomes increasingly disordered or chaotic. Kauffman (1993, 1995), in fact, describes the linkages between the binary elements in terms of the genetic phenomenon of epistasis, which is the dependency between genes in respect of their fitness contributions to the organism of which they are a part. It should be clear that the fitness contributions of genes encoding for specific traits or features depends, to a large extent, on whatever other genes are simultaneously present in the genotype. Thus a gene that encodes for dense and heavy bones may be of benefit to a large, flightless bird, but it is unlikely to be useful in the case of a small, flying one. The main point here is that the fitness contributions of genes are rarely independent of other genes; rather, genes interact in complex ways to co-constrain their respective fitness contributions to the organism of which they are a part.

[4] This profile of frequently occurring small avalanches and infrequently occurring large avalanches should bring to mind the probability distribution of belief cascades observed by Glinton et al (2010).

Epistatic linkages thus determine the extent to which genes are free to vary: if no epistatic linkages exist, then each gene is free to vary independently of every other gene; if such linkages do exist, then the fitness contribution of one gene is, in part, determined by the genes to which it is epistatically connected. In the context of Boolean networks, the complex regime emerges at intermediate levels of epistasis. Thus, if the number of epistatic connections is low, the system is highly ordered, whereas if the number of epistatic connections is high, the system is highly chaotic. Between these two regimes, at intermediate levels of epistatic connectivity, we encounter the complex regime. It is here, Kauffman (1993, 1995) argues, that we see the potential for self-organization and evolutionary change.

In formalizing the notion of epistasis, Kauffman (1993, 1995) develops a model that has come to be known as the NK model. It is essentially a model for generating evolutionary fitness landscapes with different degrees of ruggedness. Within the model, there are two parameters: N (which refers to the number of genes) and K (which refers to the number of epistatic linkages between the genes). In terms of our discussion about Boolean networks (and belief systems), N refers to the number of binary elements in the network (the number of beliefs in the belief system) and K refers to the number of linkages between the binary elements (the number of inter-dependencies between the different beliefs). Kauffman (1993, 1995) shows that by using these two parameters, we can generate a variety of 'tunable fitness landscapes', each of which differs with respect to the ruggedness of their topography. Thus, when $K = 0$, there are no epistatic connections between the elements (genes, beliefs, or whatever), and thus each is free to change independently of the others (the fitness contribution of each gene is independent of all others in the genotype). In this situation, Kauffman shows that the fitness landscape is very smooth, and the landscape contains a single global optima. For $K > 0$, the fitness landscape becomes increasingly rugged, so that when $K = N – 1$ (the maximum value of K), the landscape is maximally rugged.

The effect of K, therefore, is to determine the extent to which adjacent points on the fitness landscape are correlated – what Kauffman refers to as the correlation structure of the landscape. This is important, because it impacts on the extent to which evolutionary processes, occurring via a gradual process of genetic change, are possible. To better understand this, consider a system composed of 3 binary elements. There are (2^3) 8 possible individual variations within this system (e.g. <000>, <001>, <011>, and so on), and an adjacent individual within this system is an individual that differs with respect to a single binary value (e.g. <000> is adjacent to <001>). If we assume that each individual occupies a particular point on the fitness landscape, then the effect of K is to determine the correlation between the fitness values assigned to adjacent individuals. K, in other words, determines the structural profile of the fitness landscape against which evolutionary processes take place. High values of K produce landscapes that are like jagged moonscapes, with soaring peaks and plummeting cliffs; low values of K produce landscapes that are like the archetypal English countryside, with gentle slopes and rolling hills.

So what does this discussion of NK models and fitness landscapes have to do with our notions of cognitive consistency and socially-mediated cognitive change? Well, we have seen that one model of a belief system represents beliefs as binary variables, with each binary variable reflecting an individual agent's belief in a particular fact. We have also seen that the notion of epistasis could be applied to the problem of representing the inter-dependencies between individual beliefs. In this case, the linkages between beliefs would be cast as epistatic linkages, and the contribution of each particular belief to the overall level of cognitive consistency would depend on whatever other beliefs were possessed by the individual. By applying Kauffman's NK model to this notion of 'epistatic binary belief system', we can begin to see that each type of belief system can be associated with a particular kind of 'fitness landscape', one whose topography depends on the number of beliefs in the system (N) and the linkages between those beliefs (K). This fitness landscape will provide information about the relative fitness of different belief configurations, and the fitness of each belief configuration will be determined by the internal consistency of the belief states. What we have, therefore, is a way of representing the inter-dependencies between beliefs, and a way of representing the fitness of different belief configurations in terms of their cognitive consistency. The high points on this fitness landscape indicate states of high cognitive consistency (to which an individual will typically be attracted); the low points indicate states of low cognitive consistency (from which an individual will typically be repelled). The pursuit of more consistent cognitive states (or fitter solutions), in this case, emulates the individual's efforts to arrive at internally consistent cognitive states – the individual is constantly driven to adopt fitter (and therefore more consistent) variants of their current cognitive state (providing that such fitter states are available, or at least reachable). For each particular configuration of beliefs, we can derive a measure of fitness based on the consistency between the beliefs, and we can then use this to evaluate the relative desirability of different belief configurations. In the context of this model, we can assume that individuals will be motivated to seek out local optima within the fitness landscape – they will move from lower to higher points on the fitness landscape by adopting beliefs that are progressively more consistent. Of course, the process of belief change within this model is analogous to the process of mutation in evolutionary models: an individual changes her beliefs by random mutation of the current belief configuration (switching the value of one or more binary variables), and thereby moving to a different point on the fitness landscape.

At this point, it may help to have a concrete example of the application of the NK model in order to exemplify what has been discussed thus far. The example presented here is based (loosely) on the work of Ahouse et al (1992) and the associated commentary provided by Monge and Contractor (2003). In particular, it takes the example described by Monge and Contractor (2003) and applies it to the case of coalition plan evaluation described by Rasmussen et al (2009). As we will see in subsequent sections, the work of Ahouse et al (1992) is important because it attempts to apply Kauffman's NK model to the co-evolutionary dynamics of *multiple* belief systems. Furthermore, the introduction of belief systems that specifically target the domain of coalition planning establishes a nice linkage with recent work on the modelling of causally-connected belief networks (Rasmussen et al., 2009)[5]. As we will see, such work may provide a means for us to apply the NK model to issues of cultural change and ideological stability.

Let us assume that an individual, *i*, has a belief system consisting of three beliefs, $B1_i$, $B2_i$ and $B3_i$. These three beliefs reflect the individual's attitudes to aspects of the military coalition planning process and the relationship of these aspects to positive planning outcomes (e.g. the generation of a high quality coalition plan). These beliefs could be, for example:

1) that plans should be developed in as much detail as possible,
2) that plans should be developed as rapidly as possible, and

[5] The work relating to idea networks and belief epidemiology is also relevant here (Simpkins et al., 2010).

3) that plans should based on as much feedback as possible.

These three beliefs can each take a value of 0 or 1, with 1 reflecting strong endorsement of the belief and 0 reflecting little or no endorsement of the belief. Thus, an individual whose belief system has the form <100> would strongly believe in the importance of developing detailed plans but would not be overly concerned about the speed of plan development or the amount of feedback on which the plan was based.

The cognitive consistency of a particular cognitive state (i.e. configuration of beliefs) is determined by the inter-dependencies between the individual beliefs. These inter-dependencies may, let us assume, be of either a positive or negative nature, so that beliefs may either reinforce other beliefs (contributing to high cognitive consistency) or contradict other beliefs (contributing to low cognitive consistency). For instance, an individual's belief that plans should be developed in as much detail as possible (B1) may be difficult to reconcile with the fact that plans should be developed as quickly as possible (B2); however, it may sit very comfortably alongside the belief that plans should be based on high levels of feedback (B3). Within the NK model, these relationships correspond to epistatic ties between the beliefs and, as discussed previously, they influence the topographic structure of the fitness landscape against which we judge the fitness of specific belief configurations. For the sake of argument, let us assume that the number of epistatic ties on each belief is 2. This gives us our value of K in the NK model (N, of course, is equal to 3 because the belief system comprises 3 beliefs). The dependencies between the specific beliefs are as follows:

1. B1 contradicts B2 and reinforces B3 (if B1 = 0, then B2 = 0 and B3 = 1)
2. B2 contradicts B1 and B3 (if B2 = 1, then B1 = 0 and B3 = 0)
3. B3 contradicts B2 and reinforces B1 (if B3 = 1, then B2 = 0 and B1 = 1)

We can now define K_{pqi} as the epistatic constraint for agent *i* between belief p_i and q_i, where:

K_{pqi} = 1 if p reinforces q, and $p_i = q_i$, else 0

K_{pqi} = 1 if p contradicts q, and $p_i \neq q_i$, else 0

K_{pqi} = 0 if p neither reinforces or contradicts q

K_{pqi} = 1 if p and q are the same belief

Thus, if p were B1 and q were B2, and we were to assume that B1 contradicts B2, then we would get:

K_{pqi} = 1 if $B1_i$ contradicts $B2_i$, and $B1_i \neq B2_i$, else 0

If an individual had the belief <010>, K_{12i} would be 1 since $B1_i$ and $B2_i$ are not equal to each other ($B1_i = 0$ and $B2_i = 1$).

Now, clearly the kind of beliefs possessed by a specific individual will determine their overall level of cognitive consistency. An individual with the belief configuration <110> will have greater inconsistency (and thus more cognitive dissonance) than an individual with the belief configuration <100>, simply because the latter belief system reduces the conflict between B1 and B2. For any individual, we can define $W(B1_i)$ as the level of cognitive consistency based on an individual's belief in B1. Given that K=2 in our model, we know that the contribution of B1 to the overall measure of cognitive consistency will be based on the other two beliefs, namely B2 and B3. In general, the contribution of each belief (b) to an individual's overall level of cognitive consistency can be denoted as W(b_i), where:

$$\mathrm{W}(\mathrm{b}_i) = \Sigma(\mathrm{w}_q \mathrm{K}_{pqi}) \tag{9}$$

Thus, the contribution of each belief to the overall measure of cognitive consistency is the sum of the epistatic constraints between p and all q beliefs, weighted by a parameter (w_q), which measures the significance of that constraint.

The overall level of cognitive consistency for an individual with a specific configuration of beliefs can now be defined as:

$$\mathrm{W}_i = 1/\mathrm{N}\ \left(\Sigma(\mathrm{Wb}_i)\right) \tag{10}$$

which can be interpreted as the fitness (cognitive consistency) of a particular belief configuration.

Now that we have computed a measure of fitness in terms of cognitive consistency, we can begin to model the 'evolution' of cognitive states based on the changes in cognitive consistency that result from changes in one or more beliefs. If we consider an individual with the belief configuration <001>, then we can see that, with respect to the dependencies between various beliefs, the belief configuration <101> represents the most optimal nearest neighbour in terms of cognitive consistency. As such, we would expect an individual to endorse B1 as well as B3 because these beliefs reinforce one another. We would not expect an individual to endorse B2 and B3 because these beliefs contradict one another and thus constitute dissonant cognitions.

Once an individual has reached a local optima in terms of the consistency of their beliefs, we might expect such a cognitive state to be highly stable. We can see that once the form of an individual's belief system is <101>, it is impossible to move to a neighbouring belief configuration without reducing overall cognitive consistency. As such, the configuration <101> is a highly stable cognitive state. All else being equal, we should not expect this belief configuration to change. If this profile of belief states were to characterize the mental life of a human agent, then we might expect them to be very resistant to cognitive change; they would, in all likelihood, emerge as very 'stuck in their ways', and it would be almost impossible to persuade them to think about things in a different way (i.e. to adopt a different cognitive outlook). In fact, the only way such an individual could be persuaded to change their views would be if we could vary more than one belief at a time. In this case, we might encourage them to shift to an alternative mindset of equal consistency, namely <010>. This configuration of beliefs is equally consistent to <101>, but it requires a rather radical shift in belief states. In fact, all three beliefs need to change in order to move from one cognitive state to the other. While such transitions may seem insignificant when we are dealing with a belief system consisting of only three beliefs, the scale of the cognitive state changes becomes more profound as the number of beliefs (and dependencies between the beliefs) increases. It may, in fact, be totally unrealistic to assume that large numbers of beliefs could all change at the same time in any reasonably complex belief system (one comprising, say, in excess of 100 beliefs).

One value of the kind of model presented here is that it may help us understand more about the psycho-dynamics of cognitive change. Obviously, there is a rich psychological literature here to which we cannot begin to do justice to in the context of this paper. A couple of points are, however, worth noting about the potential for cognitive change in epistatic binary belief systems. One point is that a high number of epistatic links between beliefs may make it difficult to modify belief configurations that already occupy local optima. Such cognitive states can, in general, only be changed by making rather large-scale jumps to alternative points in the fitness landscape. If we want to encourage cognitive change in such systems, a couple of options seem worth considering (see Monge & Contractor, 2003). Firstly, we can attempt to reduce the number of epistatic connections between beliefs. This might be accomplished by changing an individual's perceptions of the linkages between beliefs or (in the case of causal knowledge) modifying their knowledge about the causal contingencies between particular states of affairs. In the case of our planning example, we might attempt to introduce new types of planning technologies that support the rapid assembly, evaluation and revision of military plans. This would act to reduce the perceived incompatibility between the three beliefs. A second strategy we might adopt to encourage cognitive change relates to the number of beliefs contributing to overall states of cognitive consistency. Thus, if we were to introduce additional beliefs that made *independent* contributions to cognitive consistency, then this would provide more freedom for beliefs B1, B2, and B3 to change without necessarily reducing the overall level of cognitive consistency for an individual agent (much, of course, depends here on the nature of the epistatic linkages between the new beliefs).

Coupled Belief Systems

Thus far in this section we have considered cognitive change as something that is driven by an individual's internal motivation to establish consistent cognitions. It should be clear, however, that individual cognition is also influenced by social factors. The network models of both Watts (2002) and Glinton et al (2010) attempt to capture this social context, and this is what makes them interesting as network-based models of collective cognitive flux. Now that we have discussed the intra-individual forces that govern cognitive change, it is time to consider the role of social forces in influencing such change. It is time, in other words, to consider the interplay between network dynamics at both the intra-individual (belief network) and inter-individual (social network) levels.

One attempt to model the influence of social factors on individual cognition is provided by Ahouse et al (1992). This study is highly relevant to the present discussion because it attempts to extend the NK model outlined above in order to accommodate the effects of social influence on individual cognition. Ahouse et al's (1992) approach is based on an extension to Kauffman's NK model, called the NK(C) model. This model was developed by Kauffman to provide insight into co-evolutionary processes. It extends the NK model with an additional parameter, C, which represents the inter-dependencies between the traits of *different* species in terms of their fitness contributions to a given individual. Thus, whereas K represents the linkages or dependencies between the genes of a particular individual, C represents the linkages or dependencies between the genes of *different* individuals. We can make sense of this notion by thinking about the fact that the relative fitness of particular traits (as manifested by a particular individual) depends, to a large extent, on the kind of traits possessed by other individuals who occupy the same ecological niche. Thus the fitness implications of a gene that encodes for large, dense bones may be relatively minor if the bird in question inhabits an environment devoid of predators. In this case, the bird can simply dispense with flight and not worry about the consequences. Things are clearly not so rosy, however, if some of the other species in the bird's environment acquire teeth, claws and a taste for flightless bird! Now, the fitness implications of a gene for large, dense bones are a cause for great concern. Whereas before, large, heavy bones were of little consequence to the bird (and may even have been of adaptive value), now they negatively affect the ability of birds to survive and reproduce. All this, of course, simply serves to remind us that no one gene is necessarily good or bad from a fitness perspective. What matters is the kind of genes that get expressed elsewhere, both within the same individual and in other individuals which (for better or worse) must be interacted with. As with many things in life, the success or failure of one's own strategies largely depends on whatever strategies are adopted by others in one's social environment.

The addition of the C parameter is thus an important addition to the NK model. It represents the fact that the fitness of any individual is often highly dependent on the features possessed by other individuals. As one individual evolves a particular set of features, it changes the fitness of features possessed by other (inter-linked) individuals. As one species adapts to the exigencies of its environment, so it changes the kind of exigencies that other species must deal with. And, as we have seen, a trait whose contribution to evolutionary fitness is irrelevant at one point in time may suddenly become highly significant as the traits of other species change. The image that emerges is thus one of a highly dynamic inter-dependence between the evolutionary fitness landscapes of different species. The evolutionary landscapes of species are coupled to one another, such that as one species moves about on its fitness landscape, seeking out locally optimal design solutions, the topographic structure of other landscapes to which that species is connected are systematically deformed. And once the individuals on the newly deformed landscape begin to adapt and explore their new terrain (again searching for local optima), their actions will feed-back to the landscape of the original species, perhaps undermining the integrity of recently scaled peaks.

We can now begin to see how the cognitive influences between agents in a social network might be modelled. Whereas the original application of the NK model (to the notion of intra-individual cognitive consistency and cognitive change) adopted the notion of a fixed fitness landscape against which individuals were motivated to seek out progressively fitter solutions, the image that emerges from a consideration of the NK(C) model is one of a highly dynamic fitness landscape that is in a constant state of flux. The source of such dynamism, in a social network context, is the inter-dependence between the cognitive states of different social actors. Thus, whether a particular belief configuration is associated with a high fitness score, and thus occupies a lofty peak on a fitness landscape, is not just a function of the consistency of one's own internal cognitions; it also depends on whatever other beliefs are adopted by individuals with whom one interacts. If the individuals within one's social network[6] adopt beliefs that are different from one's own, then one may be influenced to modify their beliefs, particularly if the fitness of one's own beliefs is less than that possessed by one's peers.

The notion of coupled belief systems is thus important because it provides a means for us to represent the influence relationships between individuals in a social network. In place of a single, individual belief system, striving for internal cognitive consistency against a fixed fitness landscape,

[6] The ties or links in such networks may be of a variety of forms, e.g. kinship ties, communication ties, reporting ties so on.

we now have multiple, dynamic closely-coupled belief systems. As one individual strives for cognitive consistency, so they influence the relative fitness against which other's cognitive states are evaluated. And whereas before we encountered individuals who were often stuck in highly stable cognitive states (i.e. local optima) and therefore resistant to social influence, we now have a situation in which the dynamics of intra-individual cognitive change are highly sensitive to the profile of cognitive state fluctuations elsewhere in the social network. As one individual changes their beliefs, so they influence the stability of cognitive states in those individuals to whom they are connected via social ties. In essence, the collective beliefs of all individuals in a social network provides the cognitive niche in which the cognitive state of any given individual evolves. As Monge and Contractor (2003) comment:

> *"...adding the C component into the NK models makes the situation much more complicated since an individual's cognitive consistency can now be reinforced or contradicted by the beliefs of other individuals in the system. In order to achieve the highest level of cognitive consistency fitness value an individual must now not only have beliefs that reinforce one another, but also have ties with other individuals whose beliefs reinforce the individual's beliefs. If, at any time, any of these other individuals change their beliefs the focal individual's cognitive consistency will be altered. Hence, unlike the NK models, the fitness landscape for each individual is no longer constant...Instead, each individual's fitness landscape is constantly changing depending on the beliefs of other individuals with which this individuals has ties and their potentially changing beliefs." (pg. 283)*

What the NK(C) model gives us, then, is a model of coupled belief systems in which the fitness landscape associated with each individual's belief system is dynamically influenced by the beliefs adopted by other (connected) individuals. The precise nature of the coupling relationships are, it should be clear, determined by the structure of the social network in which an agent is embedded. Thus any given agent will be influenced by those agents to which they are connected, and the C parameter in the NK(C) model will be determined by these connections. The result is a complex model of belief system dynamics in which the role of social network structure may be expected to interact, in as yet largely unknown ways, with the internal cognitive dynamics of specific individuals.

Stability and Change in Coupled Belief Systems

The foregoing discussion may have created the impression that cognitive states, under the influence of social forces, are forever apt to change – that cognitive stability is something that emerges only on rare occasions. This is not necessarily the case. There are situations in which globally stable cognitive states can emerge in coupled belief systems. In particular, NK(C) models may occasionally result in Nash equilibria, in which the local optima of one individual is consistent with all the local optima of the individuals to which that individual is connected. In the case of coupled belief systems, a Nash equilibrium means that each individual within a network (or sub-part thereof) has reached a local optima that cannot be improved upon, providing other individuals persist with their own specific belief configurations. Since this applies to all individuals, no individual has anything to gain by changing their beliefs, and thus the system as a whole has reached a state of global cognitive stability in which no further (internally-driven[7]) change is possible. As Monge and Contractor (2003) comment, a Nash equilibrium means:

> *"that each individual has acquired the most cognitively consistent belief configuration that can be achieved while taking into account the belief configurations of all other individuals in the system." (pg. 284)*

Although Nash equilibria may be rarely encountered in the type of network contexts we are considering here, comprised as they are of large numbers of interacting individuals[8], if they do occur, then a stable profile of belief configurations emerges in which no individual benefits by unilaterally changing his or her beliefs. Unless the system is disturbed by some exogenous force, the system will become cognitively inert, ceasing to engage in any further form of cognitive change.

As Monge and Contractor (2003) are quick to point out, Nash equilibria are not necessarily good for a system. What such equilibria mean, in the case we are considering here, is that each individual has attained the highest level of cognitive consistency that can be achieved given the profile of cognitive states adopted by all other individuals. The presence of Nash equilibria does not necessarily mean that the population, as a whole, has settled on the most collectively-optimal solution. Just as the Nash equilibria for some forms of the Prisoner's Dilemma Game may mean that each individual has to settle for strategies that result in sub-optimal payoffs for all the participants, the emergence of Nash equilibria in the case of coupled belief systems may mean that each individual has to make do with rather unsatisfactory levels of internal cognitive consistency. While this may very well reflect the mundane reality of everyday social living, with social interaction requiring all manner of personal compromises and concessions, it is nevertheless important that we understand how the dynamics of information flow and influence in networked communities of cognitive agents contributes to the emergence of these states. Moreover, it is possible that as we understand more about the relationship between network structure and cognitive dynamics, we will encounter new opportunities for influencing the long-term cognitive outcomes of such systems, biasing their behaviour in ways that benefit both the individual and the social collectives of which they are a part.

Cultural Models and Collective Cognition

Thus far, we have adopted a model in which the focus has been on individual agents and their social relationships. The belief systems in this model were deemed to be those of specific individuals, embedded in social networks. Belief systems need not, however, be modelled at the level of individual agents. An alternative approach is to develop an aggregate model of beliefs (and their associated linkages), which reflects the coarse-grained regularities associated with the belief systems of multiple individuals from a specific population. The technique of Cultural Network Analysis (CNA) is designed to produce such models (Sieck & Rasmussen, 2007; Sieck et al., 2010). CNA is a

[7] This obviously does not exclude the possibility that change may be driven by external factors; i.e. factors external to the social system under consideration.

[8] Kauffman (1993, 1995) argues that Nash equilibria rapidly emerge in NK(C) systems where $K > S * C$. Since S in this equation (at least in our case) represents the number of individuals that are interacting in the model, we can see that Nash equilibria are unlikely to emerge in any system except those in which the number of epistatic links between individual beliefs is exceptionally high. Of course, much depends here on how we interpret the parameter S. In a later section, we will encounter an alternative interpretation of S that equates it with the aggregate belief system of a specific cultural group. In this situation, the S parameter may be quite small, and the potential for Nash equilibria to emerge may be correspondingly high.

methodology that synthesizes techniques from a range of disciplines, including cognitive anthropology, cultural psychology, naturalistic decision-making, and decision analysis. Its aim is to develop models (referred to as cultural models) that capture the key (cognitive) commonalities among a group of individuals in terms of their shared concepts, beliefs and values. Cultural models reveal the relative differences and similarities between social groups, and they can be used as the basis for discriminating between specific cultural and sub-cultural groupings. An example of a cultural model, in this case applied to the domain of military planning, is shown in Figure 4. The nodes in this diagram represent concepts and properties associated with the domain of military planning, and the linkages between the concepts reflect the community's beliefs regarding the relationships and dependencies between the concepts. The links associated with a plus sign reflect a positive association between the concepts (e.g. 'Level of detail in specification of action' has a positive effect on 'ease of execution'), whereas the links associated with a minus sign reflect negative associations between the concepts (e.g. 'ability to anticipate problems' has a negative effect on the 'riskiness of a plan').

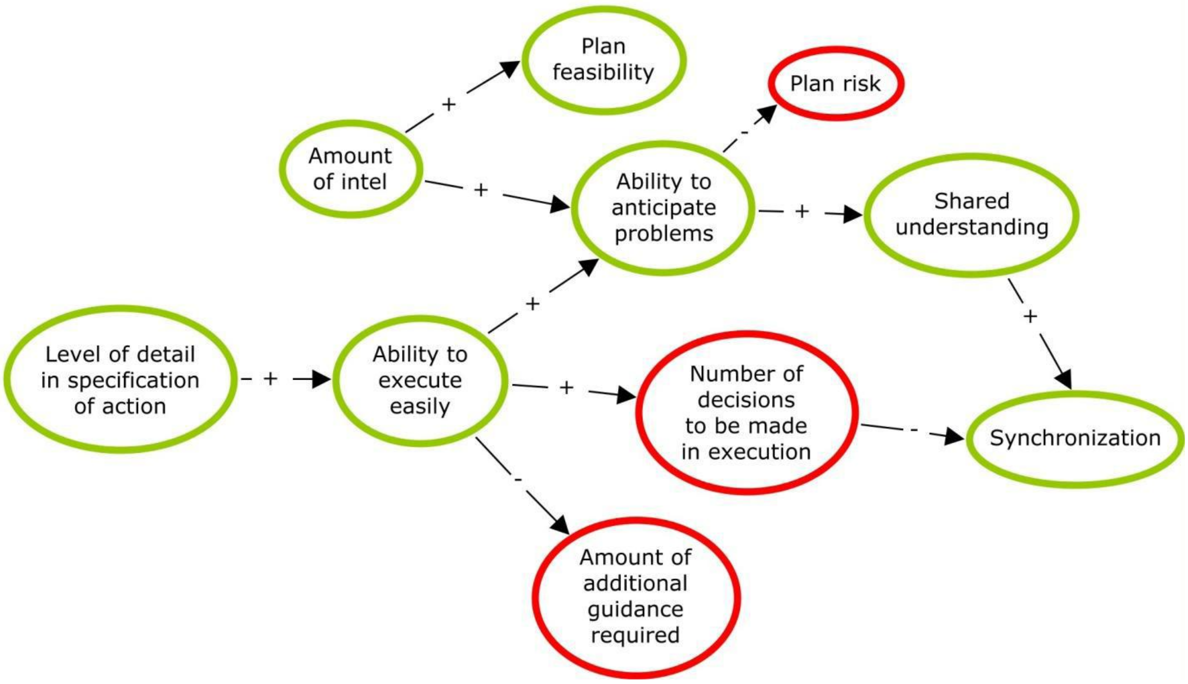

Figure 4: Cultural model showing the dependencies between concepts in the domain of military coalition planning.

What should be clear from Figure 4 is that there is a correspondence (of sorts) between the structure of a cultural model and the kind of epistatically-linked belief systems that we have been discussing thus far[9]. An important difference between the two models is that cultural models abstract away from the specific details of individual (agent-level) belief systems, focusing instead on the common features of such systems. Cultural models are created based on the inter-individual similarities between the belief systems of specific individuals, and thus, for any given population, we

[9] Thus, in order to provide a Boolean network representation of cultural models we could cast the nodes as binary variables where a value of 1 means that the associated feature has the value 'high' and a value of 0 means that the associated feature has the value 'low'. The positive and negative links between the nodes then become epistatic ties that either reinforce or contradict other nodes.

may end up with a relatively small number of models, each of which reflects a culturally-significant grouping of individuals within the larger population. The primary significance of this shift away from individual belief systems to more collective, aggregate belief systems is that it enables us to think about the potential for cognitive change mediated by a variety of cultural-level influences. Something similar to this is proposed by Ahouse et al (1992). They recognize that rather than considering the belief systems of specific human individuals, we could consider the belief systems associated with entire nation states. Thus, if we consider the general beliefs held by, say, people from the US, Russia and Western Europe, we can develop an NK(C) model in which N represents the beliefs held by the various populations, K represents the dependencies between the beliefs held by each population, and C represents the dependencies between the beliefs held by different nation states (e.g. between US and Russia). As suggested by Ahouse et al (1992), these beliefs could be about the relative merits of military spending for security, for the overall economy and for the extent of government involvement in domestic social policies. Inasmuch as these respective belief systems form a coupled belief system, then we may expect the cognitive dynamics of one population to influence the fitness landscapes associated with the populations to which it is connected. Thus, a shift in public opinion in one population may result in, perhaps unexpected, changes in other populations. Not only does this extension of the coupled belief system model highlight the potential inter-dependencies between nation states (or cultural groups) in terms of their collective cognitive outlook, it also indicates that international influences need not necessarily involve the direct manipulation or targeting of foreign populations. Instead, based on the coupled belief system model, cognitive change may be effected in foreign populations by focusing on the attitudes and opinions of domestic populations. By introducing the right profile of cognitive change in local populations[10], the fitness landscapes against which the ideologies, beliefs, attitudes and values of other nation states and cultures take shape are systematically altered. Providing we know the pattern of dependencies between the belief systems of specific nations and cultural groups, and providing the cognitive changes required by the local population are not overly adverse[11], it seems entirely plausible that cognitive changes could be inculcated in one population in order to indirectly affect the cognitive dynamics of other, connected populations.

The transition from individual to collective belief systems raises another important issue, this time concerning the issue of cross-cultural cognitive stability and the emergence of Nash equilibria. Recall that one concern raised in respect of conventional social networks (in which the actors are individual human agents) was that the number of individuals in the models made the emergence of Nash equilibria highly unlikely. Nash equilibria are, in fact, only likely to emerge in systems where the number of epistatic ties (K) outnumbers the number of individual agents multiplied by the number of dependencies between the agents' beliefs (C). Thus in cases where we are dealing with large numbers of interacting belief systems (as in most social network contexts), we are unlikely to see the emergence of Nash equilibria. In the case of cultural models, however, the number of belief systems is relatively small: it corresponds to the number of individual cultural models that feature in any given simulation. Providing the internal complexity of the cultural models is sufficiently rich (with

[10] Actually, what is important here is not so much the actual profile of beliefs adopted by a domestic population, so much as the perceptions of the target population. Thus, what is important is not actual change, so much as perceived change. The perceptions of a target population could be influenced by the way in which various national media portray the beliefs, values and attitudes of domestic populations.

[11] Although as mentioned above, actual cognitive change may be unnecessary.

large numbers of dependencies between the cultural model elements), and providing the number of dependencies between the beliefs of separate cultural models is relatively small, then we may expect Nash equilibria to be a relatively common occurrence in the context of cultural-level cognitive dynamics. This presents an interesting, although potentially perplexing, possibility: cognitive stability will tend to dominate in network simulations pitched at the cultural level (by virtue of inter-cultural forms of influence), but this will contrast with a tendency for cognitive instability in network simulations pitched at the level of social networks (by virtue of the number of belief systems present in any given social network). Thus, once we take a cultural-level perspective of belief system dynamics, the potential for highly stable, albeit collectively sub-optimal, system states becomes a realistic possibility. However, when we consider the cognitive dynamics of socially-embedded belief systems, we see little possibility for cognitive stability and cognitive flux seems to rule the day.

Of course, whether you find this state of affairs perplexing or not probably depends on your intuitions regarding the relationship between cognitive models pitched at the individual and collective levels. If you see the structure of collective-level models (e.g. cultural models) as driven primarily by the interactions between individuals, and do not see individuals as being influenced by the structure of collective models, then you will tend to favour a dynamic view of cultural models in which inter-cultural forces make little contribution to the cognitive dynamics of the global belief system (the entire collection of belief systems at the individual and collective levels). If, on the other hand, you see a potential role for collective-level models in influencing cognitive change at the individual level (perhaps because of the influence exerted by cultural artefacts, rituals and symbols), then you will tend to see collective-level models as moving towards a Nash equilibrium and the potential for individual cognitive change as strongly influenced by forces operating at the cultural, national and societal levels.

A final point about cultural models is worth mentioning here. It is that the content of such models often reflects the causal knowledge that a group of individuals has about some domain. Thus the linkages between the elements of a cultural model often reflect knowledge about cause and effect relationships, such as the knowledge that high levels of shared understanding contribute to improved inter-agent synchronization in the case of US military planners (Rasmussen et al., 2009; Sieck et al., 2010) (see Figure 4). This serves to remind us that the content of a particular belief system is quite open; it could consist of attitudes towards particular things, or it could consist of the cause and effect relationships governing the transitions between a set of system states. In the former case, our measure of cognitive consistency closely approximates conventional notions of cognitive dissonance, with the perceived incompatibility between different attitudes (e.g. a belief in animal rights and a belief that is ok to wear fur products) contributing to cognitive inconsistency and belief change (at least in some cases). In the latter case, our measure of cognitive consistency could be grounded in the overall causal coherence of the belief states, with violations of domain-relevant causal knowledge contributing to our measure of cognitive inconsistency and providing the impetus for cognitive change. To help us get a better grip on this, consider what we said at the very beginning of this sub-section in relation to the processing of situation-relevant information:

> *"...consider a situation where a human decision-maker is attempting to formulate a comprehensive picture of a situation based on a body of incomplete and uncertain information. Suppose that the individual has already received information that leads her to form certain beliefs. Now, whenever new information about the situation is received,*

> *the information will be evaluated with respect to those pre-existing beliefs. In cases where a conflict is encountered (i.e. new information contradicts a prior belief), the individual will either have to revise her pre-existing beliefs or discount the newly received information. The point here is simply that whenever an agent entertains multiple beliefs, their background (e.g. causal) knowledge often enables them to detect inconsistencies between belief states, and such inconsistencies can influence decisions about whether new information is accepted or rejected. In cases where conflicting information is accepted, the agent will be forced to revise existing beliefs (or at least downgrade the certainty assigned to those beliefs); and in cases where such information is rejected, we can begin the understand its lack of influence in terms of an existing complex of inter-dependent belief states."*

So we have come full circle. From a consideration of existing models of belief propagation, we have encountered a model that sees the dynamic profile of collective cognition as emerging in response to a variety of psycho-dynamic and socio-cultural influences. Firstly, we saw that individuals might come under internal pressure to change their profile of beliefs based on the perception of specific inconsistencies and conflicts. In this case, cognitive change is motivated by a need for consistency between beliefs, and the individual strives to minimize inconsistent cognitions. Secondly, we saw that cognitive change could come about as a result of forces that operate at the social level of analysis. In this case, the individual is influenced by the beliefs possessed by other individuals, and cognitive change is likely to strike a balance between social influence and cognitive dissonance. Finally, we saw that there might exist collective forces that influence the cognitive dynamics of individuals within specific cultural groupings. This level of analysis emphasizes the potential role that inter-cultural relationships play in stabilizing the cognitions of individual group members. As yet, we have very little understanding of how these various levels interact to shape and constrain the profile of collective cognition. Most studies limit their attention to one particular level of analysis; however, it seems likely that multi-level models and simulations will be required to reveal the true nature of real-world collective cognitive flux. The NK(C) model explored in this section, with its grounding in complexity science and the study of self-organized systems, provides a useful starting point for multi-level simulations. In addition, the use of networks to model information flow and influence at all levels, establishes a sensible, and potentially productive, linkage with the emerging science of networks (Barabasi, 2002; Buchanan, 2002; Watts, 2003). Together, we can hope that such approaches will enable us to gain a better understanding of the various forces and factors that shape the profile of cognition at both the individual and collective levels.

Network Structure and Collective Problem Solving

When thinking about the mental life of human agents, it is common to make a distinction between cognitive states and cognitive processes. Cognitive states, like states of belief, are the aspects of our mental life that we typically recognize as the intermediate products of some thought process. Cognitive processes, on the other hand, are the processes that govern the transition between cognitive states. Cognitive processes are typically seen to involve some form of information processing; i.e. the transformation and manipulation of information-bearing structures. And it is cognitive processes, like perception, reasoning, thinking, and the comprehension of language, which have tended to dominate much of the work in cognitive psychology.

When it comes to collective cognition, we can continue to make a distinction between cognitive states and cognitive processes. Cognitive states, like states of belief, have clearly been the focus of the paper thus far. We have thus talked about cognitive states as corresponding to the specific configuration of beliefs adopted by particular agents, and we have attempted to provide some insight into how such states may be influenced by forces acting at the individual, social and even the cultural level. The current section aims to extend the discussion to a consideration of cognitive processes in collective (multi-agent or group) situations. But before we embark on this discussion, we need to consider what it is that makes a process cognitive, and whether it makes sense to talk of cognitive processes as being distributed across the elements of a social network.

There are, in fact, two possible interpretations of what collective cognitive processing might mean. Firstly, we might see collective cognitive processing as consisting of those processes that influence the cognition of individual agents. On this view, cognitive processes are the things that go on inside the heads of individual agents, and the purpose of multi-agent simulations and group-level analyses is to understand how these inner processes are affected by social forces. On the other hand, we might see collective cognitive processing as reflecting a commitment to something like distributed cognition (Hutchins, 1995). In this case, cognitive processes would be seen as distributed across the elements of a social network, with social actors making the same sort of contribution to cognitive processing as the neuronal elements of the biological brain make to the cognitive processes of the individual agent. In the case of distributed cognition, the interaction of the agents within a social network constitutes the basis for cognitive processing, and cognitive processes are seen as the properties of the larger system in which agents participate.

The notion that cognitive processes might be distributed across the elements of a social network, in the same way as they are (in the more conventional case) distributed across a nexus of neuronal (and perhaps extra-neuronal (see Smart et al., 2010)) elements, is clearly compatible with the claims of the distributed cognition movement. However, not everyone is happy to concede that cognitive processing *is* something that can take place outside the heads of individual human agents. Harnad and Dror (2006) thus argue:

> *"…cognition takes place entirely within the brains of cognizers…The causes and effects stretch more distally, but not the cognition; cognition begins and ends at the cognizer's sensor and effector surfaces."*

Clearly, what is needed in order to resolve this dispute is some way of determining what makes a process cognitive. This would enable us to determine whether the information processing that takes place at a collective, multi-agent level is really something that should be seen as a legitimate form of cognitive processing. Unfortunately, there has been very little progress in our understanding of what constitutes a cognitive process. The main problem is that cognitive processes are often defined by ostension. We can therefore point to specific examples of cognitive processing (e.g. perceiving, reasoning, thinking and so on), but establishing what it is that makes something a legitimate member of the class of cognitive processes is a much more difficult undertaking. Adams and Aizawa (2001, 2008) favour a view of cognition that highlights the role of representations with 'intrinsic' as opposed to 'derived' intentionality. Unfortunately, it is not entirely clear what is meant by the notion of 'intrinsic intentionality', or when we confront representations whose content is intrinsically given. Rowlands (2006) offers a definition of a cognitive process as:

> *"one that: (i) is required for the accomplishing of a cognitive task, (ii) involves information processing, and (iii) is of the sort that is capable a yielding a cognitive state" (pg. 32).*

In this definition, the notion of a 'cognitive task' is defined by ostension, and the notion of a cognitive state is construed as a genuinely representational state; i.e. a state that can be seen as representational in virtue of its satisfaction of a host of additional criteria. The main problem here, of course, concerns the fact that we are still relying on ostensive definitions for the notion of a cognitive task. We also encounter problems with respect to the precise conditions under which a physical state should count as one that is 'genuinely representational'. Perhaps, as Clark (2010) suggests, the best that we can do in this case is to define a cognitive process in terms of its behaviour-supporting role. That is, a cognitive process should be defined with respect to the kinds of behaviours that it makes possible:

> *"What makes a process cognitive, it seems to me, is that it supports genuinely intelligent behaviour....To identify cognitive processes as those processes, however many and varied, that support intelligent behaviour may be the best we can do."* (Clark, 2010, pg. 92-93)

Of course, this definition leaves the notion of what constitutes 'intelligent behaviour' uncomfortably vague, and it also fails to address the issue of whether cognitive processes can be socially distributed. Thus, if we see intelligent behaviour as the behaviour of a specific individual, then it is not clear whether the actions of other individuals should really count as part of the processing that makes such behaviour possible. Smart et al (2010), for example, talk about the way in which scientific open access initiatives[12], in conjunction with global information networks, serve to facilitate creative insight and intellectual progress in the domain of scientific endeavour. As Harnad (1999) rightly notes, systems like the World Wide Web enable us to accomplish something akin to 'scholarly skywriting' – scientific theories, thoughts, ideas, experimental results, and sometimes data, are made available in ways that are increasingly accessible to fellow academics and scientific colleagues. Should the collective actions of the scientific community, in this case, count as part of the cognitive processing substrate that makes possible the pursuit of productive shifts in individual-level thinking?

For the purposes of the current paper, we propose to view collective cognitive processing as a form of distributed cognition. That is, we see cognitive processes as things that can be distributed across the elements of a social network, and we see interacting agents as potentially contributing to the collective realization of these processes. This position enables us to see the coordinated behaviour of multiple agents as constituting a form of cognitive processing; it enables us, in essence, to see social networks as part of the material fabric that makes collective cognition possible.

One study which has attempted to examine collective cognitive processing in network contexts is Mason et al (2005). Unlike many of the studies discussed so far, Mason et al (2005) used human subjects rather than synthetic agents. They attempted to examine the ability of groups of networked human subjects to collectively explore a problem space and find optimal solutions within that space. The subjects had to guess a number between 0 and 100, and they were awarded points based on a

[12] http://www.eprints.org/openaccess/

score associated with that number. A continuous 'fitness' function determined the score associated with each number, but this function was not made available to subjects. Instead, on any given round of the experiment, subjects had to choose a number based on the feedback they received from both their own guesses and the guesses of their immediate neighbours. Thus, in a totally connected network (see Figure 5), each subject could see the scores associated with the guesses of all other participants, and they could (if they so wished) choose to imitate successful neighbours by copying their guesses on subsequent rounds. Collective problem-solving performance was assessed by calculating the average score of subjects on each round of the experiment. Thus, in order to be successful, subjects had to explore the structure of the fitness landscape, as provided by the fitness function, and rapidly converge on the most optimal solution (i.e. number) available. The cognitive process in question is thus a form of collective problem-solving in which the search efforts of multiple individuals are pooled to create a measure of collective cognitive success (i.e. the ability of the group to find an optimal solution).

In order to assess the effect of network structure on problem-solving performance, Mason et al (2005) organized the human subjects into communication networks with different structural topologies (see Figure 5). As just mentioned, in the totally connected network condition, each participant could see the solutions proposed by all other participants as well as the fitness score associated with their solutions; in other network conditions, each participant could only see the solutions and scores of a more limited number of individuals, namely those individuals to which they were directly connected. The question we can now ask is how will the nature of the communication network influence the collective problem-solving behaviour (and performance) of the participants? How, in other words, will network structure affect the performance profile of collective problem-solving?

What Mason et al (2005) found was that in conditions where there was a single local optima (i.e. a fitness landscape with a single peak), the network structures that supported the most rapid dissemination of information were the most successful. Thus, when the fitness landscape had a single peak, subjects tended to converge more quickly on the global maximum in the totally connected, small world and random network topology conditions. This result is explained in terms of the speed at which information is spread through the network. Thus, in the single peaked condition, each solution with a higher score provides information about the best direction for future search efforts (all routes uphill lead to the same optimal solution), and, as such, it is helpful to have a situation whereby information about the results of collective search efforts are distributed as widely as possible throughout the problem-solving community.

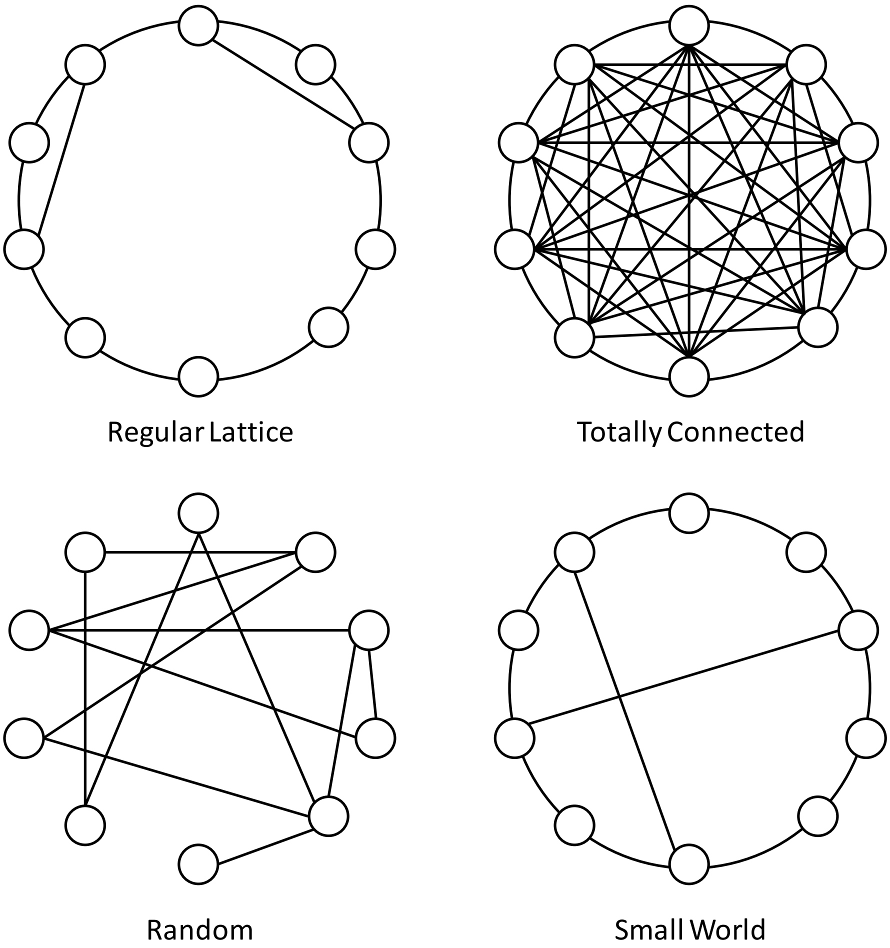

Figure 5: Examples of the network structures used by Mason et al (2005).

Things are very different, however, when the fitness landscape has a more rugged, multi-modal structure; i.e. when there are multiple local optima (and one global optima). In this case, we might expect rapid information dissemination to once again result in rapid convergence on a particular solution, but whether this solution is the best one available depends largely on how lucky the participants are with their initial guesses. If the participants refrain from converging too quickly on a particular solution, they might discover a more globally-optimal solution, and it is for this reason that we might expect network structures that limit the rate of information dissemination to benefit collective problem-solving performance.

This is pretty much what Mason et al (2005) discovered. They found that when a multi-peaked solution landscape was used, participants found the global solution most quickly in the small-world network condition. The topology of the small-world network seemed to provide just the right amount of social influence for optimal performance – it supported a certain amount of independent exploration, but it did not inhibit rapid convergence on optimal solutions. As Mason et al (2005) comment:

> *"In the small-world network...the participants were segregated into different spatial regions, but the information could travel quickly through 'short-cuts', allowing for different locally-connected groups to explore various regions of the problem space. Thus, while one locally connected group might latch onto a local maximum, the small-world*

> *topology decreases the probability that everyone will follow their lead before another sub-group finds the global maximum. Once any subgroup finds the global maximum, the information can spread quickly to other subgroups...”*

Results similar to Mason et al (2005) have also been reported by Lazer and Friedman (2006, 2007), this time using experiments involving synthetic agents. As with Mason et al (2005), Lazer and Friedman (2006, 2007) also examined the effect of network structure on collective problem-solving performance using a search task. However, in this case, the search task consisted of a search for optimal solutions using the NK model paradigm discussed earlier in this paper. A specific NK model (N=20, K=5) was used to generate a fitness landscape, and agents were required to explore the landscape by generating variants of the 20 bit design solution. At the beginning of each simulation, agents were randomly assigned locations on the fitness landscape, and at each round of the simulation they could explore the local terrain of the landscape by randomly switching one of the binary variables associated with their current solution. This modification to a single binary variable provided information about the relative fitness of neighbouring solutions and thus indicated whether fitter alternatives were available. In addition to receiving information about the fitness of their own solutions, agents also received information about the fitness of solutions as proposed by their network neighbours (i.e. the agents to whom they were directly connected). If a neighbouring agent proposed a solution that was fitter than either the agent’s current solution or their modified solution, then the agent adopted the solution of its neighbour.

Using this procedure, Lazer and Friedman (2006, 2007) discovered a trade-off between what they called exploration and exploitation. Exploration here is the tendency of agents to explore the solution space independently of other agents, and exploitation is the tendency of agents to adopt the solutions proposed by other agents. The result of this trade-off is a profile of topology-dependent performance in which networks with low average path lengths (e.g. totally connected networks) yield better performance in the short-term (compared to networks with higher average path lengths (e.g., linear networks)), but worse performance in the longer term (again compared to networks with higher average path lengths). We can make sense of this pattern of results if we think of the agents as a bit like explorers parachuted into a mountainous landscape at night, equipped with radios, flashlights and altimeters. Imagine that the goal of each explorer is to find the highest peak in the landscape, and thus all explorers begin to move uphill as soon as they touchdown. As they move ever upward, the explorers can communicate with each other using their radios, reporting the results of their altimeter readings. And if agents hear that one of their colleagues has found a higher point on the landscape, then they all converge on that location and begin searching from that particular point. Networks that support full two-way radio communication between all the explorers (i.e. totally connected networks) deliver a profile of good short-term, but poor long-term performance: all agents rapidly converge on the highest location found at the outset of the search, but given the rugged nature of the terrain the highest initial location is not necessarily the one that leads directly to the highest peak. If one agent lands on the top of a small hill, and all the other agents land at the base of a separate, taller hill, then all agents will move to the top of the small hill and remain there; they will fail to discover the most optimal solution (the peak of the tall hill) because the agent atop the small hill will communicate a higher altimeter reading, and all agents will immediately converge on that location. If the rate at which explorers can exchange information is reduced, as is the case with alternative network topologies, then our intrepid explorers can avoid the distracting influence of other agents and potentially discover heights that are much loftier than those that might otherwise have been discovered.

The results of studies exploring the effect of network structure on collective search tasks thus points to a common conclusion: different types of network topology can affect the rate at which information propagates within a problem-solving community, and this can compromise a group's ability to discover globally-optimal, long-term solutions. When the network topology supports rapid rates of information transfer, individuals may be inclined to settle on sub-optimal solutions on the basis of initial shared information. Such results may be of considerable significance when it comes to understanding a range of phenomena such as groupthink (Janis, 1982), production blocking (Diehl & Stroebe, 1987) and the common knowledge effect (Stasser & Titus, 1985)[13], all of which seem to be characterized by a group's inability to find optimal solutions based on some form of precipitant interaction or early information sharing.

Of course, the implications of Mason et al's (2005) and Lazer and Friedman's (2007) findings should not be over-stated. The results are important in terms of contributing to our understanding of the relationship between network structure and collective problem-solving; however, the conclusions we can make from such research findings are limited in a number of ways. Firstly, the results do not reveal that a particular pattern of network connectivity will always deliver the same profile of performance, even when the nature of the problem-solving tasks resembles that used in the studies. The reason for this is that the primary determinant of collective performance seems to be the rate at which information is disseminated, coupled with the tendency of agents to always copy superior solutions. This particular profile of rapid information dissemination and high social influence may be seen in some problem-solving contexts, but, in the real world, technological and social factors will often conspire to undermine the actual rate of information flow and influence through a social network. Thus, inter-agent influence may be undermined if superior solutions cannot be copied because of copyright laws or patents. Moreover, just because a network structure supports rapid information dissemination this does not mean that the actual flow of information through the network must be necessarily rapid. Agents or nodes within the network can effectively modulate the speed at which information is transmitted by selectively ignoring, or by only intermittently processing, information. In human social networks, there are clearly a variety of factors that might contribute to the rate of information distribution. These include things such as the tendency to hoard information, willingness to cooperate, vulnerability to copying/transmission errors[14], and the level of trust between neighbouring actors. Also, of course, in situations involving mobile ad hoc

[13] Hinsz, Tindale and Vollrath (1997) have also highlighted some of the dangers associated with a group's over-reliance on shared information. Such insights, in combination with the results reported here, should give us pause for thought when it comes to notions of shared situation awareness (Nofi, 2000) and shared understanding (Smart et al., 2009a). Inasmuch as the interventions used to enhance shared situation awareness and shared understanding depend on the sharing of common sets of information, it is important that we do not create a situation in which group-level problem-solving abilities are undermined as a result of the attempt to achieve some other human factors objective.

[14] Lazer & Friedman (2007) evaluated the impact of copying errors in their computer simulation studies. They report that, in the long-run, systems with high error rates in the copying process outperformed those in which copying errors were minimized. The explanation for these results seems to be the same as that proposed for the effect of network structure on performance, namely that "Error rates in copying….alter the balance between exploration and exploitation in the system, increasing the amount of experimentation but reducing the rate with which successful strategies spread" (Lazer & Friedman, 2007).

networks (MANETS), many nodes may be expected to have only occasional or intermittent connectivity, and this may impede the effective rate of information spread through the network[15].

A second limitation of the studies reviewed in this section concerns the nature of the problem-solving task used to explore collective problem-solving. It should be clear that the search tasks used by Mason et al (2005) and Lazer and Friedman (2007) are only one type of task that could be used to study collective problem-solving. Many types of real world collaborative problem-solving involve significant amounts of specialization in which agents work on particular parts of a problem and then attempt to coordinate their activities with respect to some larger, overarching problem-solving objective. It seems unlikely that the results of Mason et al (2005) and Lazer and Friedman (2007) can generalize to these more differentiated and hierarchically-structured tasks, because such tasks feature complex inter-dependencies between the behaviours of particular agents. In the case of the search tasks, the success of one particular agent has no effect on the success of any other agent, and thus the solutions proposed by one agent are independent of all the others. However, some of the most interesting cases of real-world collaborative problem-solving, such as military planning, involve situations where the suitability of candidate solutions as proposed by one agent are heavily dependent on the solutions proposed by other agents. In cases like hierarchical military planning, we therefore seem to confront a situation that is more akin to the notion of the coupled fitness landscapes discussed in relation to the NK(C) model. In this case, the topographic structure of the fitness landscape for one part of the problem is systematically deformed whenever solutions for other parts of the problem are proposed. The upshot is that it seems unclear whether the conclusions of Mason et al (2005) and Lazer and Friedman (2007), made in the context of collective search tasks, can be extended to other types of collective problem-solving. Further work needs to be undertaken in order to examine this issue.

A final concern regarding the studies reviewed in this section concerns the nature of the putative cognitive processes that are being explored. Inasmuch as we can say that the studies are exploring a process that is distinctly cognitive in nature, albeit one that is obviously distributed across multiple agents, then it is clear that we are dealing with a cognitive process that is unlike, say, memory, reasoning or perception. To say that the studies of Mason et al (2005) and Lazer and Friedman (2007) provide some general insight into the nature of network-enabled socio-cognitive processing, we therefore need to be able to extend the work to other types of collective cognitive processing. One example here might be to look at the effect of network structure on distributed memory processes. A number of studies have sought to examine memory in a social context, looking at the role of social influence and interaction on the processes of encoding, recall and contagion (see Barnier et al., 2008, for a review). Few studies, however, have systematically explored the effect of communication network structure on the social distribution of memory processes. Given the clear role that networks play in constraining the opportunities for inter-agent interaction and influence,

[15] This should give us pause for thought when it comes to considering the apparent limitations of MANETS. Rather than see intermittent connectivity as an unfortunate side-effect of the current capabilities of wireless technology, something to be eliminated by future design efforts, the current set of results casts the connectivity profile of MANETS in a more positive light. If intermittent or periodic connectivity effectively retards the rate at which information is transmitted, then it seems entirely possible that, at least in some contexts, the connectivity profile of MANETS may have adaptive value. Rather than exerting a uniformly negative effect on collective cognition, connectivity limitations may sometimes play an important role in helping a community of problem-solving agents come to a high quality cognitive outcome (for example, a particular decision or an accurate shared interpretation of situation-relevant information).

we might expect to see a number of interesting results once we think about collective memory (and other types of collective cognitive processing) from a network-oriented perspective.

Exploring the Effect of Dynamic Networks on Collective Problem Solving

Network structures that promote the rapid dissemination of information have, we saw, a worrying tendency to impair collective cognitive performance, in at least some problem-solving contexts. In particular, as the structure of the solution landscape becomes progressively more rugged, and the complexity of the problem to be solved becomes correlatively more complex, the rate of information dissemination seems to become increasingly important in determining the final outcome of collective processing. Networks that facilitate the rapid transfer of information between the elements of a social network can, it seems, result in precipitant forms of information sharing which result in premature convergence on a sub-optimal solution.

One way of addressing this problem, of course, is to enforce structural change in the social network, modifying the network topology in ways that effectively reduces the rate at which information is disseminated. However, rapid information dissemination is not always a bad thing. In some cases, as when the solution landscape is relatively smooth and there is a single global optimum, then rapid, widespread communication may actually benefit collective problem-solving performance. Arguably, what is required is some way of dynamically adapting the network structure to suit the nature of the problem-solving task confronting a community of agents. Instead of searching for a network structure that is uniformly beneficial for collective problem-solving, perhaps we need to think more about dynamic networks, networks that dynamically and adaptively modify their connectivity in response to the demands of specific problems. In fact, there are compelling reasons to think that a more adaptive and dynamic approach to structural change might benefit collective cognitive performance. As we will see in subsequent sections, networks may best support cognition (at both the individual and collective level) by a process of what we refer to as 'adaptive coupling'; i.e. the temporally-specific coupling of various resources into a flexibly-configured and dynamically-bounded cognitive system[16].

Many studies that aim to investigate the patterns of information flow and influence in social networks do so using networks with fixed, static topologies. There is no reason, however, to assume that such networks necessarily exhaust the space of cognitively-interesting network simulations, or even that such networks necessarily represent the kind of networks typically encountered in cases of much real-world cognitive processing. As such, in this section, we report the results of a preliminary attempt to evaluate the contribution of dynamic networks (networks whose structural and functional topology changes throughout the course of cognitive processing) to collective cognition. The specific aim of the study was to investigate the impact of constructive changes to a network (realized by the progressive addition of links) on collective problem-solving performance. The study relied on the same problem-solving paradigm as that used by Lazer and Friedman (2007), and the NK models were of the same complexity; i.e. all simulations used NK models with parameters of N=20 and K=5. These parameters yield fitness landscapes that are moderately rugged, with a few hundred local optima and high correlations between proximate solutions. As Lazer and Friedman (2007)

[16] The notion of adaptive coupling thus goes hand-in-hand with the notion of cognitive extension (Clark, 2008), according to which non-biological elements can sometimes constitute part of the mechanistic substrate of a cognitive system.

comment, these kind of problem spaces probably best capture "the essence of most interesting problems that individuals and organizations in the real world face—rugged, but not chaotic" (pg. 674). Before running the simulations, 1000 different NK models were generated using the aforementioned parameters. These 1000 models were then used in experimental simulations in which a collection of agents (population=100) were tasked with the exploration of the problem space and the discovery of optimal solutions. The experiment involved the manipulation of two independent variables: Network Growth Rate (NGR) (which represents the rate at which links are added to the network) and Network Growth Delay Period (NGDP) (which represents the period that must elapse before the first link is added to the network). In total there were 7 levels of the NGR variable and 6 levels of the NGDP variable, giving a total of (7*6) 42 experimental conditions. Within each condition, the 100 agents were tested against the same set of 1000 NK models (i.e. the same set of 1000 NK models were used for each condition), resulting in a total of 1000 simulations for each condition (i.e. a total of 1000*42 = 42,000 simulations). At the beginning of each simulation, none of the agents were connected (i.e. there were no links between any of the agents), but throughout the course of each simulation links were added in order to connect the agents together into networks of increasing density (see Figure 6). As is indicated by Figure 6, the networks in this study were generated in a particular way. Links were added to the agent community by first selecting an agent at random from the total population of 100 agents. A second agent was then selected (again at random) from the subset of agents that were already connected to at least one other agent (i.e. the agent possessed at least one link). The two agents were then connected via the addition of a link[17]. The network growth law in this study is thus a form of preferential attachment (of unconnected nodes to connected ones), but there is no bias towards connecting to nodes that have a particular number of links (e.g. the most links)[18].

Each simulation was completed when at least one of the following termination conditions was encountered:

1. 1500 processing cycles had elapsed, or
2. all possible links had been added to the network and the network was thus fully connected, or
3. all the agents were connected into a single network component, they had all converged on a common solution, and at least 20 processing cycles had elapsed with no new solutions being generated by *any* of the agents[19].

The rate at which links were added to the emerging network depended on the value of the NGR variable for the specific experimental condition, and the cycle at which the first link (or set of links) was added depended on the value of the NGDP variable. Thus, if the NGR variable was set to 1.0 and

[17] This is why, in Figure 6, we see the emergence of a single network component after the addition of only 100 links.

[18] Clearly, there a number of ways in which the agent network could be generated. These alternative growth laws are the subject of ongoing investigations.

[19] On each processing cycle, an agent selects one element of the 20-bit solution string with which it is associated at random. This means that over the course of a 20-cycle period an individual agent may fail to completely explore all neighbouring solutions (i.e. they may fail to select some elements of the 20-bit solution string). However, because condition 3 is based on the fact that we are considering the solutions generated by *all* the agents (and all the agents have converged on a common solution), then the chance that the agents will collectively fail to explore all neighbouring solutions is very small (i.e. 2.8 E-45).

the NGDP variable was set to 50, then the first link would be added on the 50th cycle of the simulation, and 1 link would be added to the 'network'[20] every cycle thereafter. The levels of the NGR variable used in this study were 0.1 (1 link added every 10 cycles), 0.2 (1 link added every 5 cycles), 0.5 (1 link added every 2 cycles), 1 (1 link added every cycle), 2 (2 links added every cycle), 5 (5 links added every cycle) and 10 (10 links added every cycle). The levels of the NGDP variable were 0 (no delay period), 10 (a delay period of 10 cycles), 20 (a delay period of 20 cycles), 30 (a delay period of 30 cycles), 40 (a delay period of 40 cycles) and 50 (a delay period of 50 cycles).

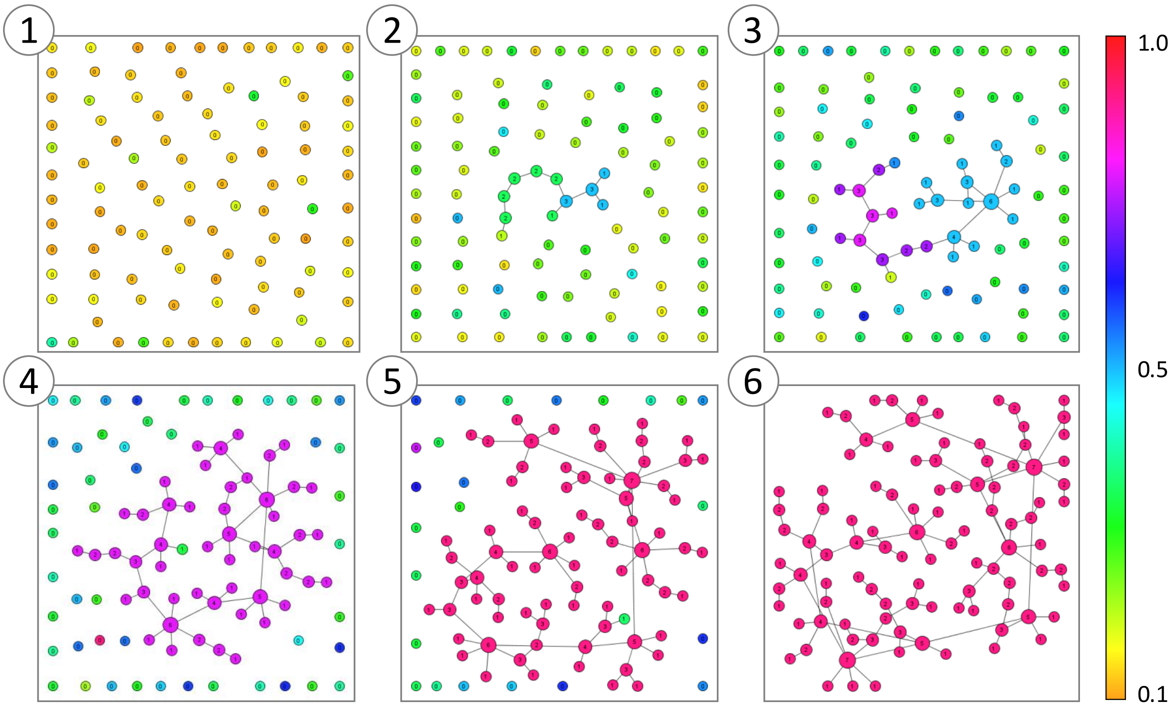

Figure 6: Figure showing the progressive addition of links during the early stages of a simulation. The graphics numbered 1 to 6 show the 100 agents (represented as nodes) at the outset of the simulation (tile 1) when none of the agents are connected, and at the 100th processing cycle (tile 6) when all the agents are connected into a single network component. The color coding of the nodes in these tiles indicates the value of the solution associated with each agent, with orange/yellow values indicating poor solutions and purple/red values indicating good solutions. As can be seen from the figure, the agents in the simulation gradually discover progressively better solutions throughout the course of the simulation.

At the start of each simulation, each of the 100 agents was assigned to a particular solution within the NK model space selected for the simulation. Agents were assigned the same value for each NK model, thus the starting point for each agent in each NK model across all experimental conditions was the same. The simulation then consisted of a number of processing cycles during which the following steps were applied to each agent (X):

1) the current solution (S) for X was compared to all the solutions of X's immediate neighbours in the network (if they had any)

[20] Obviously, the term 'network' is something of a misnomer here. Until sufficient links have been added, the system consists mostly of disconnected agents.

2) if the solution of a neighbouring agent was found to be better than S, then X adopted the solution of its neighbour
3) if none of the solutions of X's neighbours were better than S, then X modified S (by randomly selecting and flipping one of the 20 binary values comprising S) to generate a new solution (S')
4) if S' was better than S, then X adopted S' as its current solution; otherwise, X adopted S as its current solution

Each simulation was run until one of the aforementioned termination conditions had been reached.

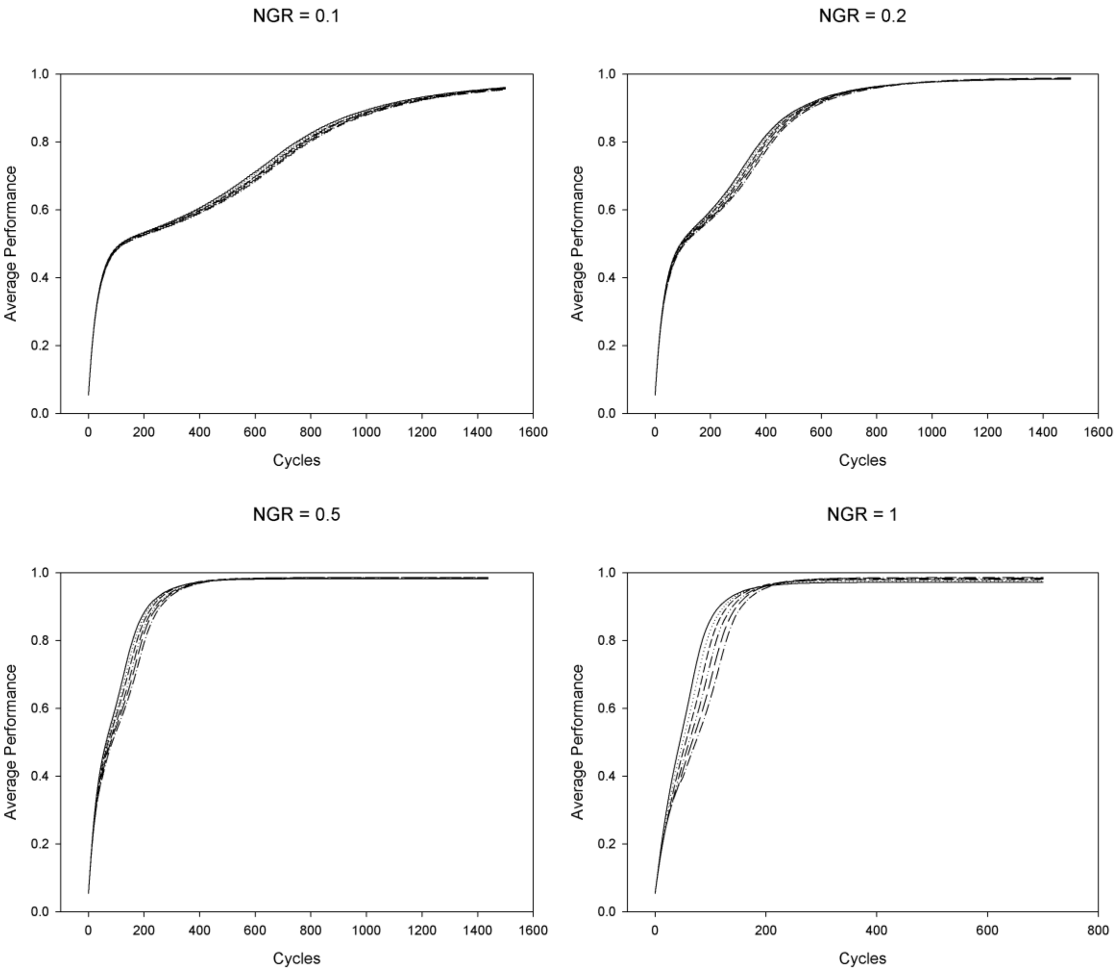

Figure 7: Average performance over 1000 simulations in networks with growth rates of 0.1 (1 link added every 10th cycle), 0.2 (1 link added every 5 cycles), 0.5 (1 link added every 2 cycles) and 1 (1 link added every cycle). The lines on each chart show the effect of the initial delay period, with longer delays causing a rightward shift in the performance curves (see Figure 8 for a legend).

Figure 7 and Figure 8 summarize the results obtained from this experiment. Both figures show the average performance of agents over the course of successive processing cycles for different levels of the NGR and NGDP variables. The average performance is calculated as the average score associated with the solutions adopted by all agents averaged across all the agents within a particular cycle (100 data points) and across all the simulations within a particular treatment condition (1000 data points). Figure 7 shows the results obtained for the first four levels of the NGR variable (i.e. 0.1 – 1), while Figure 8 shows the results obtained for the remaining levels (i.e. 2, 5 and 10). What these figures appear to show is that as the rate of network growth increases (i.e. as the rate at which links are added to the network increases), the length of the simulation decreases (i.e. the number of cycles that must elapse before one of the termination conditions is encountered is reduced). However, as the rate at which links are added is increased, the final performance of the agents (the quality of the final solution) also becomes more variable, with performance apparently negatively affected in conditions involving both high growth rates and shorter delay intervals. The effect of the initial delay interval, therefore, seems to be that it offsets a growth rate-related decline in collective problem-solving performance. Furthermore, at each level of the NGR variable, the effect of increasing the initial delay period seems to be a shift in the performance curve to the right, particularly during the middle part of the simulation. Thus for cycles in the middle of a simulation, the effect of the delay interval is to reduce the quality of the solutions found by agents; however, by the end of the simulation this performance deficit is eliminated. In fact, for the higher growth rate conditions, the slight performance deficit seen during the middle of the simulation is reversed at the conclusion of the simulation, with higher levels of performance being seen in the conditions involving longer initial delay periods.

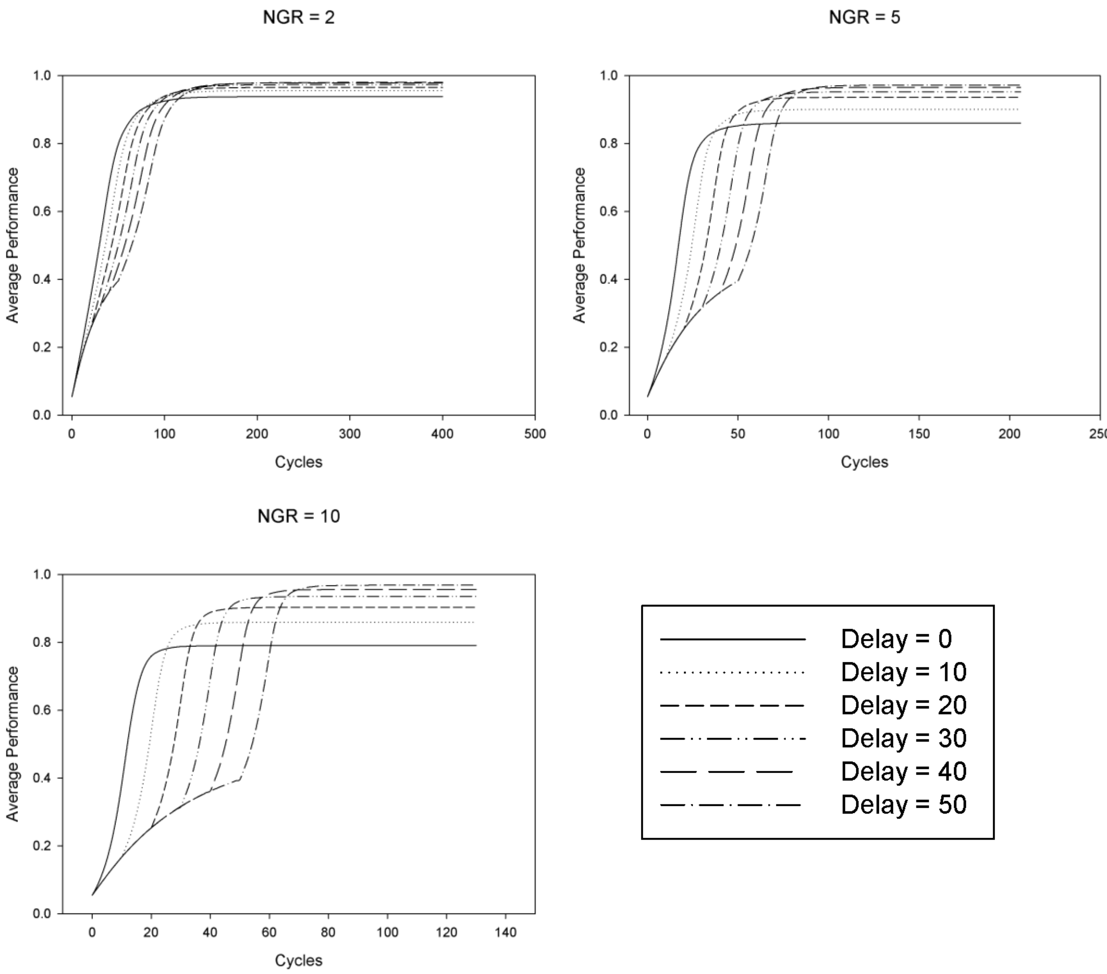

Figure 8: Average performance over 1000 simulations in networks with growth rates of 2 (2 links added every cycle), 5 (5 links added every cycle) and 10 (10 links added every cycle). The lines on each chart show the effect of the initial delay period, with longer delay periods progressively shifting the performance curve to the right.

These conclusions are backed up by the statistical analysis of the simulation data. Firstly, a two-way (7*6) between subjects factorial ANOVA on the performance scores associated with solutions at the end of the simulations revealed a significant interaction between the NGR and NGDP variables ($F_{(30, 41958)} = 148.629$, $P < 0.001$) and significant main effects for both the NGR ($F_{(6, 41958)} = 1218.754$, $P < 0.001$) and NGDP variables ($F_{(5, 41958)} = 538.756$, $P < 0.001$). Figure 9 illustrates the average final performance score obtained in the various experimental conditions. As can be seen from the figure, an increase in NGR results in a progressive deterioration of performance, and this deterioration seems to be most pronounced for conditions involving shorter initial delay periods. An analysis of simple main effects at the levels of the NGR variable reveals that differences between the various delay period conditions begin to emerge at growth rates of 0.5 links per cycle and above (i.e. NGR = 0.5). Thus the performance differences between NGDP conditions at the 0.5 level of the NGR variable were statistically significant ($F_{(5, 41958)} = 3.091$, $P < 0.01$), but below this level (i.e. growth rates of less than 0.5 links per cycle) there were no significant differences between the NGDP conditions.

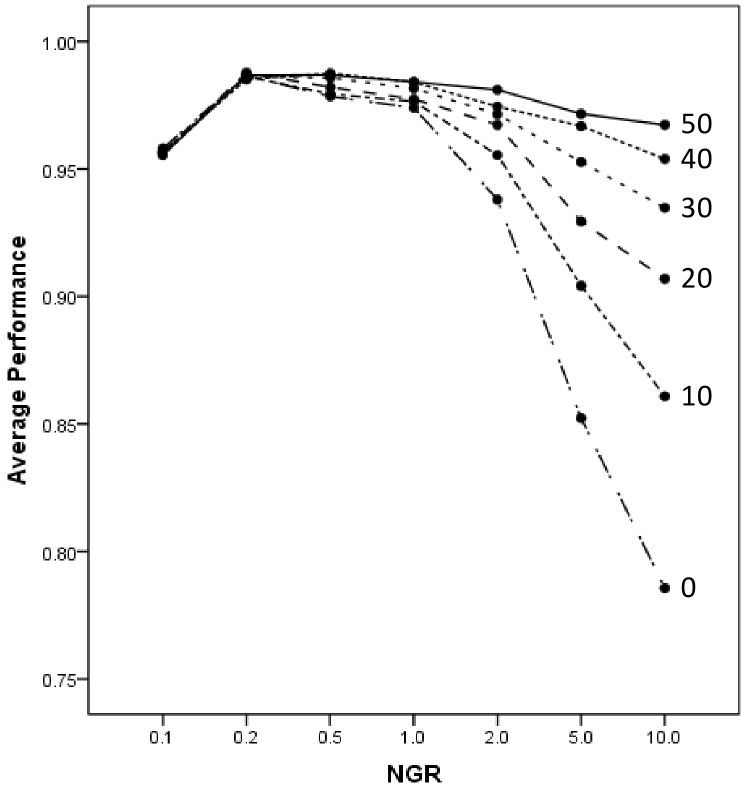

Figure 9: Average performance scores for all experimental conditions at the end of the simulations. The numbers next to the lines on the chart represent the levels of the NGDP variable. Thus the line labeled '0' represents the data obtained in conditions with an initial delay period of zero (i.e. NGDP = 0).

Besides the ability of higher initial delay periods to attenuate a growth rate-related decline in final performance scores, Figure 9 also suggests that performance was negatively affected by the slowest rate of network growth rate (i.e. NGR = 0.1). A one-way ANOVA comparing the performance data for various levels of the NGR variable revealed a significant difference between the NGR group means ($F_{(7, 41993)} = 962213.491$, $P < 0.001$), and post hoc comparisons using Tukey's Honestly Significant Difference (HSD) test revealed that performance was worse at the 0.1 level of the NGR variable when compared to intermediate levels of the NGR variable (i.e. growth rates between 0.2 and 2.0). The reason for this performance deficit at very slow rates of network growth is probably attributable to the termination conditions established for the current study. One of the termination conditions, recall, was about the maximum number of processing cycles that could be run in a single simulation. At very slow rates of growth it appears that the agent community did not have sufficient time to reach the same levels of performance as seen at intermediate rates of growth. This interpretation is supported by an inspection of the performance data for the NGR = 0.1 condition (see Figure 7). From Figure 7 we can see that by the 1500th cycle, the performance of agents had not quite reached the levels that we see in the NGR = 0.2 and NGR = 0.5 conditions. The relatively poor performance of agents in the slowest growth rate condition therefore appears to be an artefact of the specific termination conditions used for the current study.

In order to compare the performance of dynamic and static networks, we ran an additional series of simulations involving static networks. A single network was generated by adding links to a population of 100 agents until a single network component had formed[21]. The performance of agents within this static network was then assessed by running 1000 simulations using the aforementioned NK models. The results of this additional manipulation are shown in Figure 10. As can be seen from the figure, the agents settle on a stable solution very quickly – after about only 15 cycles. This contrasts with the results seen in dynamic networks where it typically took much longer for the dynamic networks to settle on a common solution. The most notable result from this manipulation, however, concerns the final performance of the agents. In the fixed network condition, the final performance of the agents reached a value considerably below that seen with dynamic networks, even when compared with the highest growth rate condition (i.e. 10 links added every cycle)[22]. The results of a one-way between subjects ANOVA on the performance data for the static network condition and the performance data obtained in the highest NGR condition (i.e. NGR = 10) revealed a significant difference between the various conditions ($F_{(7,\ 6993)} = 48537.010$, $P < 0.001$). Post hoc comparisons between the conditions using Tukey's HSD test revealed that all dynamic networks outperformed the static network (see Figure 11). We therefore arrive at the conclusion that problem-solving performance is enhanced in networks with dynamic, incremental topologies, relative to networks with fixed, static topologies.

[21] Because the method used for generating the static network was the same as that used to generate networks in the dynamic conditions, a single network component emerged only after the addition of 100 links.

[22] The average performance of the dynamic networks used in this study is also greater than the performance of the single component random networks seen in Lazer and Friedman (2007).

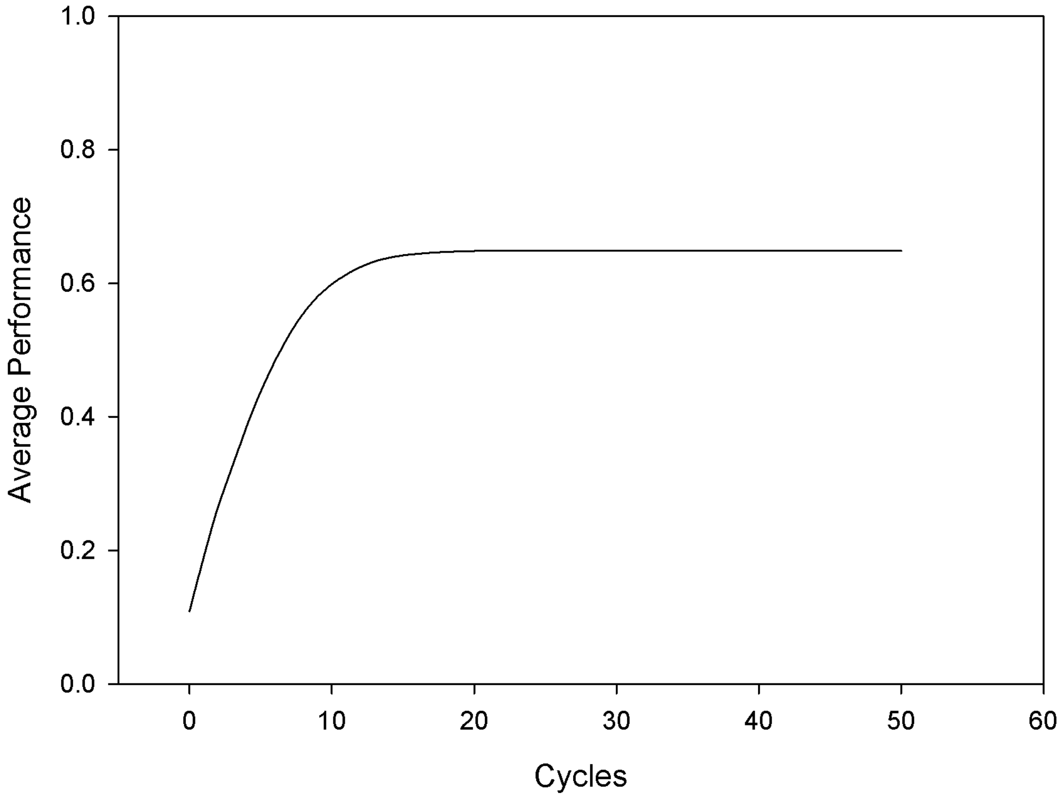

Figure 10: Average performance over 1000 simulations in a network with a fixed topology.

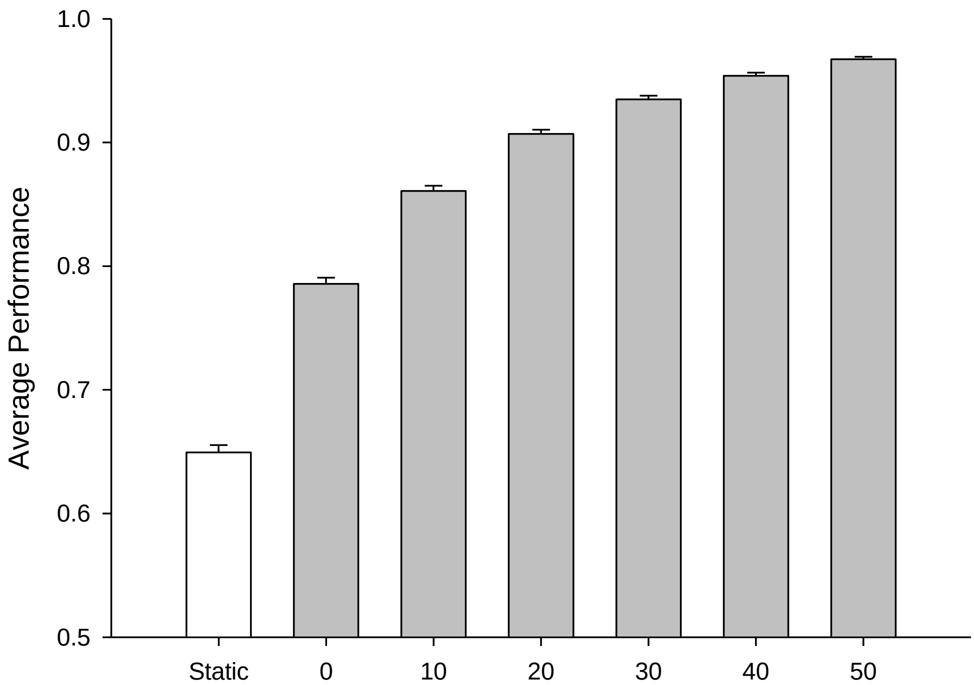

Figure 11: Graph showing the final performance scores associated with static and dynamic networks. The shaded bars represent conditions featuring dynamic network conditions at various levels of the NGDP variable (i.e. delay periods of 0, 10, 20, 30, 40 and 50 cycles). All dynamic networks had a growth rate of 10 (i.e. 10 links were added every processing cycle. Post hoc comparisons revealed that all dynamic network conditions outperformed the static network condition ($P < 0.001$).

Accounting for this particular pattern of results requires us to think about the effect of time-variant changes in network architecture on the opportunities that agents have for independent exploration of the solution space. Given what we know about the way in which agents are influenced by the superior solutions of connected neighbours, it becomes apparent that the more time agents have to independently explore the solution space, the more likely the community is, as a whole, to discover the global optimum. We have seen that when agents are connected together in a totally connected network, the rate of information dissemination is at its highest, and thus the agents are inclined to prematurely settle on a sub-optimal solution (they settle on the most optimal solution found on the first processing cycle) (see Lazer & Friedman, 2007). This explains why, in the case of dynamic networks, agents are more likely to prematurely settle on a globally sub-optimal solution in the high growth rate conditions: the presence of more and more links across successive processing cycles progressively increases the rate at which information is transmitted between the agents. It also explains why, in general, we see better performance in dynamic networks than we do in static networks (at least static networks that are created using the method described here). In dynamic networks, the extent of inter-agent influence is limited; every agent can only influence those agents to which they are directly or indirectly connected, and, at least initially, all agents begin with minimal influence (i.e. no two agents are connected). This poverty of influence, relative to the situation with fully connected networks, means that each agent has time to undertake a local exploration of the solution space before reporting the results of this local search to all other agents (when the fully connected network eventually emerges). Incrementally constructive networks therefore strike a productive balance between autonomy and influence; they give each agent the freedom to search for locally-optimal solutions without sacrificing the (eventual) benefits of collective search.

Two additional phenomena from the aforementioned study now require explanation. One is the tendency for increasing initial delay periods to negatively affect performance at the mid-stage of the simulation (recall that the performance curve is shifted to the right in most of the high delay period conditions (see, for example Figure 8)); the other is the tendency for increases in the initial delay period to counteract the deleterious effect of high growth rates on the quality of final solutions. The former phenomenon probably stems from the fact that with high initial delay periods agents have greater initial autonomy in exploring the solution space. This means they can exhaustively search for locally-optimal solutions without being influenced by the search results of other agents. Unfortunately, however, this freedom comes at a cost, because the *average* performance of a set of disconnected agents will always be worse than the *average* performance of a fully connected network of agents. When agents are connected, they can share information about the best solution currently on offer, which means that all agents can converge on the best solution. When agents are disconnected, each can only progress as far as the nearest local optimum. So even if one agent is lucky enough to find the global optimum, the *average* performance of the community will still be relatively low.

In respect of the second phenomenon (the fact that longer initial delay periods attenuate a growth-related decline in performance), the longer initial delay period supports the greater initial autonomy of agents and thus enables them to explore more of the solution space before they converge on a common solution. When no delay period exists, agents quickly become interconnected (particularly when the growth rate is very high) with the result that they are inclined to prematurely settle on a solution that, in all likelihood, is sub-optimal relative to the kind of solution that they could have found if they had been allowed to exhaustively search their local part of the solution space.

The Cognitive Virtues of Dynamic Networks

The empirical results presented in the previous section seem to indicate that, when it comes to collective cognitive processing, dynamic networks can sometimes outperform their more static counterparts. Clearly these results are, in some sense, specific to the kind of task in which the agents are engaged, and we should not necessarily expect the same pattern of results to be obtained in other kinds of problem-solving context. Nevertheless, there are reasons to suspect that dynamic networks may benefit cognition (more on which below), and this, at the very least, suggests that a comparative analysis between static and dynamic networks is something worth pursuing in the context of future research.

Dynamic networks, it is fair to say, have been somewhat neglected by those interested in network-enabled cognition. Most of the studies reviewed in this paper focus solely on networks with static topologies, but this emphasis is arguably at odds with what we see in many real world contexts. In the real world, the structure of the network topology is not fixed; rather, it changes throughout the course of information processing, perhaps in response to information processing itself, or as a response to natural processes of growth and decay. In the case of socially-distributed cognition (where the problem-solving elements correspond to social actors), the structure of the communication network may vary across many temporal scales. Only in the case of very transient episodes of information processing does it make sense to talk of networks in which the opportunities for information flow and influence between the various agents are fixed for the entire course of the problem-solving process. In most real-world contexts, agents will (sometimes repeatedly) engage and then disengage from a communication network, or they may change their profile of connectivity as new social ties are formed and old ones wither away. Even in situations where the structural topology of a network seems largely fixed, this does not mean that the 'functional' or 'effective' topology of the network is not wildly various. Thus in the case of neural information processing what seems to be important is not so much the relatively static hodological profile of specific neural circuits, so much as the dynamic patterns of neural activity which such circuits make possible. The distribution of various neurotransmitters and neuromodulators throughout such circuits seems to enable, on occasion, the effective 'rewiring' or 'reconfiguration' of the circuits in response to specific information processing challenges (Marder & Bucher, 2007; Selverston, 1995). Thus even in situations where we encounter networks with largely static topologies, this does not mean that the 'effective' structure (the structure that realizes the information processing capabilities of the focal network) is not, in some sense, dynamic.

At present, our understanding of the effects of dynamic changes in network topology on the cognitive processing potential of a network system is somewhat limited. However, there are a number of reasons to suspect that such effects may be significant. By way of introducing this idea, consider the attempt by Jerry Elman (1993) to train an artificial neural network to process complex bodies of linguistic information. Elman (1993) was interested in whether a particular type of neural network, called a recurrent neural network (see Elman, 1990), could be trained to learn about aspects of grammatical structure, such as the ability to learn about verb agreement and clause embedding in the sentence 'The girls whom the teacher has picked for the play which will be produced next month practice every afternoon' (example from Elman, 1993). The recurrent neural network architecture (see Figure 12) combines a standard three-layer feed-forward neural network with a set of context units, which serve as a kind of local memory for the network. The context units copy the activation pattern of the hidden unit layer and feed it back to the input layer alongside the next set of inputs. The result is that recurrent networks can be sensitized to temporal and sequential dependencies in the training data, learning to predict, for example, the correct linguistic category of missing words when presented with partially complete sentences.

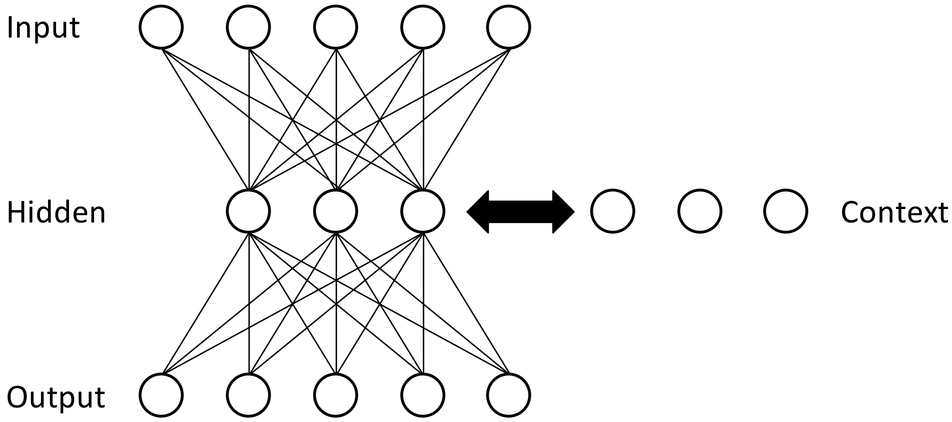

Figure 12: An example recurrent neural network architecture. In each processing cycle, the activation values of the hidden units are copied to a separate layer of context units. These values are then fed back to the hidden unit layer at the next processing cycle.

What Elman (1993) aimed to do was examine whether a recurrent neural network could learn about the grammatical structure of complex sentences, such as those exhibiting multiple clause embedding and long distance dependencies. In particular, the neural network was trained to take one word of a sentence at a time and predict what the next word in the sentence might be. As Elman (1993) comments, this task "forces the network to develop internal representations which encode the relevant grammatical information" (pg. 5).

Alas, Elman's efforts were in vain. The network completely failed to learn about the grammatical structure of the complex sentences. Not only did the network fail to develop a fully generalizable performance profile, it also failed to adequately master the data on which it was trained. In trying to account for these results, Elman (1993) tried an alternative training regime in which the network was presented initially with examples of very simple sentences and then exposed progressively to more complex ones. The aim was to isolate the point at which the network's performance broke down. At what level of sentential complexity would the network prove incapable of making further progress?

The results from this alternative training regimen were surprising. Elman (1993) discovered that when presented with staged training inputs (each increasing in complexity) the network was able to realize its original training objectives. Thus, what seems to be important to a network's ultimate ability to learn about grammatically complex sentences is that its training is structured in such a way that it is able to learn about the simple cases first. Once the network is proficient in handling these these simple cases, it can learn to cope with progressively more complex ones. It is almost as if the network's initial success with the simple cases lays the groundwork for subsequent learning about the more complex ones. As Elman (1993) rightly notes, this is an interesting result, in part because it seems to align itself well with the developmental profile of young children:

> *"Children do not begin by mastering the adult language in all its complexity. rather, they begin with the simplest of structures, and build incrementally until they achieve the adult language."*

By themselves, of course, these results seem hardly relevant to our discussion of dynamic networks; however, Elman (1993) went on to explore the effects of a further manipulation that built on the results of the staged input case. Rather than impose restrictions on the sequential order of exposure to training cases, Elman (1993) used an incremental memory solution in which the recurrent feedback provided by the layer of context units was gradually increased as training progressed[23]. The effect of this manipulation was to limit the temporal window in which linguistic inputs could be processed. It thus forced the network to focus (at least initially) on the simplest training cases. And as the memory provided by the recurrent units was increased over the course of training, so the network was able to deal with progressively more complex inputs. The effect of the incremental memory solution was thus the same as that achieved by staged training; it promoted an initial undersampling of the training data in such a way that the network's long-term ability to learn about complex regularities was enhanced. As Elman (1993) notes, this is an important discovery because it may shed light on the functional significance of a developmental progression in human cognitive capabilities. Rather than see the working memory limitations of a young children (see Kail, 1984) as a computational shortcoming that needs to be overcome in order to reveal the functional profile of adult cognition, Elman's (1993) findings suggest that such 'limitations' may play an important (and perhaps indispensable) role in children's cognitive development:

> *"Seen in this light, the early limitations on memory capacity assume a more positive character. One might have predicted that the more powerful the network, the greater its ability to learn a complex domain. However, this appears not always to be the case. If the domain is of sufficient complexity, and if there are abundant false solutions, then the opportunities for failure are great. What is required is some way to artificially constrain the solution space to just that region which contains the true solution. The initial memory limitations fulfil this role; they act as a filter on the input, and focus learning on just that subset of facts which lay the foundation for future success."* (pg. 9-10)

Elman's (1993) findings thus encourage us to think about the functional significance of what might be referred to as a network's ability to dynamically change the resources that are allocated to particular problem-solving processes. The notion of resources here is wide ranging; it can refer either to aspects of network architecture (e.g. number of nodes, link topology, etc) or to features of network processing (e.g. bandwidth, response latency, resistance to activation, and so on). Relative to networks in which such resources are largely fixed and pre-specified, such 'dynamic networks' may have capabilities that differ in interesting and surprising ways. Indeed, work on neural networks suggests that learning can be facilitated by allowing networks to dynamically reconfigure their nodes or acquire additional nodes (Ash, 1989; Fahlman & Lebiere, 1990; Schultz & Schmidt, 1991). In addition, techniques that reduce network size have been shown to have a positive impact on network performance, at least in some training contexts. Thus the progressive elimination of nodes

[23] It should be clear that this manipulation involves the use of dynamic networks. At earlier stages of the training process, the recurrent units are only active for the first few processing cycles of any trial, and they are disabled once a certain number of processing cycles is reached. This means that the functional topology of the network is effectively changed during the course of sentence processing. The deactivation of the recurrent inputs corresponds to a case where the connections between input and context units are effectively removed.

from the hidden unit layer of neural networks forces a neural network to progressively learn a more concise characterization of the regularities underlying successful performance in a given problem domain (Mozer & Smolensky, 1989a; Mozer & Smolensky, 1989b). Such results may shed light on the potential significance of regressive neural processes in which an initial excess of neurological resources are progressively pruned over the course of neurodevelopmental timeframes (e.g. Edelman, 1987)[24].

Dynamic Networks and Adaptive Coupling

One of the reasons dynamic networks benefit cognitive processing is because they enable the time-dependent adaptive coupling of individual agents into larger networked ensembles. We thus encounter what might be called the thesis of adaptive coupling:

> *Adaptive Coupling Thesis: In situations where cognitive outcomes depend on the coordinated activity of multiple resources, cognitive performance will benefit from the ability to dynamically and flexibly couple those resources into transient networks of information flow and influence. Dynamic networks support the realization of multiple time-variant patterns of functional connectivity, and these enable the component resources to adaptively coordinate their activity at critical junctures in a collective problem-solving process.*

The critical element of this thesis is the temporal dependency of the network's connectivity dynamics. Thus, the root cause of the problems in Lazer and Friedman's study is not that agents shared too much or too little information, it was that information sharing took place at the wrong time, relative to the features of the specific problem that was being solved. When the structure of the solution landscape is relatively smooth, it benefits agents to rapidly exchange information as widely as possible; however when the structure of the solution landscape is chaotic, it benefits agents to engage in a period of initial exploration followed by subsequent pooling of search results. The key point here is that what counts is the temporal profile of inter-agent information exchange – the adaptive, time-variant coupling of agents into highly configurable nexuses of information flow and influence.

As a means of further exemplifying the importance of time-variant adaptive coupling in collective problem-solving performances, consider the case of collective problem-solving described by the cognitive anthropologist, Edwin Hutchins (1995). Hutchins (1995) investigated the collective problem-solving capabilities of networked communities of simple agents. The agents in Hutchins' study were simple neural networks, known as constraint satisfaction networks. A constraint satisfaction network comprises a number of processing units linked by either inhibitory or excitatory connections. The processing units in Hutchins' study coded for specific environmental features, such as 'is a dog', 'barks', 'has fur' and so on, and the units were wired up in such a way that consistent features were connected via excitatory links and inconsistent features were connected via inhibitory

[24] The neural selectionist account of Edelman (1987) maintains that an initial activity-independent overproduction of synaptic resources is followed by a phase of activity-dependent synaptic selection and pruning. An alternative view is proposed by Quartz and Sejnowkski (1997) (see also Quartz, 1999). They argue for a neural constructivist account in which neural development is driven by an activity-dependent progressive elaboration of neural circuits. In either case, what is important for present purposes is the way in which the features of mature, adult cognition are progressively realized by dynamic, activity-dependent changes in neural architecture.

links. Thus the units coding for 'is a dog' and 'barks' would be connected via excitatory links, while the units coding for 'is a dog' and 'meows' would be connected via inhibitory links.

As Hutchins points out, once a constraint satisfaction network has been setup it shows properties akin to the psychological phenomenon of confirmation bias, which is the tendency to ignore or discount evidence that contradicts some initial interpretation of a situation. Thus imagine that a particular unit in the constraint satisfaction network is activated, say the unit that codes for the feature 'barks'. Once this unit is activated it will excite some units, such as those coding for 'is a dog', and inhibit others, such as those coding for 'meows'. The result will be that the network will settle into a consistent 'interpretation' of the input data. All the dog-related units will become active, reflecting the agent's interpretation of the input as reflecting the presence of a dog. Importantly, once a stable pattern of activity has been established by the network, it can be very hard for alternative interpretations to emerge. Thus if we subsequently present the network with a cat-related feature (e.g. we activate the unit coding for the feature 'meows'), what we find is that the pattern of activity corresponding to the 'dog' interpretation continues to predominate. The network has seemingly discounted or ignored evidence that contradicted its initial interpretation of the input data.

In order to explore the effect of inter-agent communication on this behaviour, Hutchins allowed the constraint satisfaction networks to communicate the results of their activity to other constraint satisfaction networks. What he found was that if the individual agents were allowed to communicate with one another from the outset of the simulation, then extreme levels of confirmation bias arose. This occurred because each agent, under the influence of information provided by other agents in the social network, was under pressure to discover a *shared interpretation* of the input data. In other words, the community of agents strove to find a set of activation patterns that satisfied the internal constraints established by inter-agent communication. The result was that agents often failed to give due weight to the evidence provided by external input data, and thus, more often than not, the community of agents tended to show more extreme forms of confirmation bias than was the case with isolated individuals.

The results of Hutchins' study thus begins to make contact with what we have seen in other studies, such as the performance profile of agents in the totally connected network simulations of Lazer and Friedman (2007). In situations where agents are allowed to exchange information with all other agents from the outset of a problem-solving process, the result seems to be a premature convergence on sub-optimal solutions. When inter-agent communication is limited (as in the case of linear network configurations or with the dynamic network simulations reported above), collective performance typically improves. The pattern of results seen in the studies by Hutchins and Lazer and Friedman are thus comparable, although clearly the problem-solving processes confronting the agents are somewhat different (the need to find optimal design configurations in the case of the Lazer and Friedman study versus the need to accurately interpret ambiguous environmental information in the case of Hutchins' study). In particular, Hutchins found that restrictions in the level of initial communication between the agents allowed each agent to establish its own independent interpretation of the environmental data, and this resulted in a reduced level of overall confirmation bias when communication was subsequently re-established. In essence, when agents were allowed to 'make up their own mind' as to the most appropriate interpretation of the input data, then the group as a whole was subsequently able to come to a more accurate shared interpretation of the external state-of-affairs. What we see, therefore, is an interesting parallel with the earlier studies on collective problem-solving. In both cases it seems that there is a delicate and temporally fine-tuned balance between initial autonomy and subsequent social influence. When agents are allowed to operate independently at the beginning of a problem-solving process, and then later allowed to communicate, the result is often a much better profile of collective performance than if extensive communication had been permitted from the very outset of the problem-solving process. As mentioned above these results are potentially relevant to a range of phenomena seen in group problem-solving contexts, e.g. groupthink (Janis, 1982), production blocking (Diehl & Stroebe, 1987) and the common knowledge effect (Stasser & Titus, 1985). All these phenomena seem to result from precipitant forms of information sharing that are enabled by highly efficient communication structures.

Aside from the parallels with Lazer and Friedman's work, there are also parallels with the work reported at the very outset of this paper concerning belief propagation in heterogeneous agent communities (Glinton et al., 2010). The task confronting agents in Hutchins' study, recall, is to establish an accurate interpretation of some external state-of-affairs based on conflicting and inconsistent environmental information. Inasmuch as we are prepared to see stable patterns of network activity in the constraint satisfaction networks as reflecting a particular profile of beliefs about the external situation, then the goals of Hutchins' agents are not that dissimilar from those in Glinton et al's study. We also see some correspondence in terms of the empirical findings of the two studies (although clearly much more work needs to be done here). In particular, Glinton et al (2010) observed that by manipulating the extent of social influence between the agents, inaccurate belief states could sometimes propagate throughout the entire multi-agent system:

> *"...it is sometimes the case that significant portions of the team come to have either no strong belief or the wrong belief despite overwhelming sensor data to the contrary. This is due to the occasional reinforcement of a small amount of incorrect sensor data from neighbors, echoing until correct information is ignored."* (Glinton et al., 2010)

The degree of influence between the agents in Glinton et al's study thus occasionally inclines the entire ensemble of interacting agents to form an inaccurate interpretation of the information they are presented with. The same conclusion applies in the case of Hutchins' study, although in this case it is unclear under what conditions the system exhibits the properties of a SOC system.

It is no doubt tempting, at this point, to come to two conclusions about the notion of adaptive coupling on the basis of the evidence presented thus far. The first conclusion is that the adaptive coupling process always involves elements that might, in some sense, be seen as internal to the network. By this we mean that the elements to be coupled are always those that comprise the focal network – neurons in the case of neural networks and social agents in the case of social networks. The second conclusion concerns the temporal profile of the coupling process. In this case, the tendency is probably to see adaptive coupling as characterized by initially low levels of coupling (e.g. limited connectivity or influence), followed by progressively higher levels of coupling at later stages. Both these conclusions are, we suggest, mistaken.

As evidence against the first conclusion, recall again the study of Elman (1993) involving recurrent neural networks and the processing of grammatically complex sentences. Recall that Elman was able to demonstrate a case of path-dependent learning for artificial neural networks. He was able to show that success in the domain of complex grammar learning could only be achieved by constraining the sequential order of training cases. Thus simple regularities had to be learned before more complex ones; attempts to deal with complex regularities from the outset resulted in the network failing to achieve satisfactory levels of performance. Although this case of grammar learning might initially appear entirely distinct from the cases of collective problem solving we have reviewed elsewhere in this paper, there is a sense in which multiple cases of poor performance have a common cause. Thus, during training, Elman's network is attempting to find an optimal configuration of weighting coefficients that will enable it to achieve the desired input/output mapping specified by the training regime. Exposure to the complete set of training cases (comprising both simple and complex sentences) results in the network establishing a set of weights that effectively puts it in the wrong part of the solution landscape, a place from which it is subsequently unable to recover. The early exposure to the complex cases thus undermines the ability of the network to achieve superior levels of performance in the longer-term; by exposing itself to early forms of informational influence (as dictated by the complex cases), the performance profile of the network, as a whole, is undermined.

The precise nature of the adaptive coupling in the case of Elman's network thus begins to emerge. The coupling is not, as we might expect, between the individual elements of the neural network; rather, it is between the network and a body of *external* training data to which the network is exposed. When a specific kind of informational contact is established between the network and the training data (one that determines the relative influence of simple and complex grammatical forms on learning), the network is able to bias its initial explorations of the weight space in such a way that it is subsequently able to accommodate the more complex cases from the training set. The right kind of informational contact, in this case, can be established in a number of ways: structuring the training data (with simple cases preceding more complex ones), dynamically changing the network's functional topology during training (e.g. deactivating the context units at various points in time), and degrading the quality of the training inputs via the introduction of noisy data. This latter manipulation, the one relating to noisy input data, may seem like an odd way to establish the right kind of informational contact between a neural network and its training environment, but the effect is just the same as the other two manipulations. The effect is, in essence, to reduce the *effective* set of training data to that which is most suited to initial learning. Elman (1993) describes the introduction of noisy data as a means of preventing the network from prematurely settling on sub-optimal weighting solutions – ones that effectively prevent the network from forming appropriate generalizations later on in training:

> *The network's learning capacity is greatest at early stages, but this is also the time when its training data are most limited, and so the network runs the risk of committing itself to the wrong generalization. If the initial data are corrupted by noise, on the other hand, the increased variability may retard learning and keep the network in a state of flux until it has enough data to make reasonable approximations at the true generalization. (pg. 17)*

This quotation reveals the subtle correspondence between the case of neural network learning and the multiple cases of collective problem-solving reviewed above. In all cases, what seems to emerge is that initial, inappropriate forms of information flow and influence (unrestricted inter-agent communication in the case of Hutchins (1995), full network connectivity in the case of Lazer and

Friedman (2007), and exposure to both simple and complex sentences in the case of Elman (1993)) are followed by the emergence of particular (sub-optimal) solutions (initial shared interpretations, initial design solutions, and initial weight settings) that subsequently preclude the discovery of more optimal solutions at a later point in time (more accurate interpretations, fitter solutions, more appropriate weight settings).

The main point to note from this discussion is that adaptive coupling is not something that need be restricted to the *internal* elements of a network; it can also apply to the *external* sources of influence to which the network is perhaps only occasionally 'connected'. What is important is not the nature of the resources that are actually being coupled together, but the way in which those resources participate in the information processing dynamics of a particular cognitive system.

So much for the view of adaptive coupling as something restricted to the internal elements of a network. What about the view that adaptive coupling involves a specific temporal profile of information flow and influence? One reason to doubt the validity of this view comes from a study by Clowes and Morse (2005), which aimed to shed light on the cognitive benefits of self-directed speech. Clowes and Morse (2005) performed simulations in which agents (implemented as simple recurrent neural networks) had to move geometric figures to various on-screen locations following the presentation of a particular command. The commands were presented to the network using a set of specialized 'word' units, and the output of these word units (what we might call linguistic information) could be fed back to the input layer of the network (via a dedicated re-entrant loop) in order to guide subsequent behaviour.

Clowes and Morse (2005) used a genetic algorithm to train groups of agents under the following conditions:

1) A control condition in which the dedicated re-entrant loop was disabled. In this condition, agents were periodically presented with external commands, but there was no internal recycling of the linguistic inputs.
2) A condition in which the word re-entrance loop was continuously active. In this condition, the linguistic inputs were recycled back to the input layer at every processing cycle.
3) A condition in which the word re-entrance loop was contingently enabled based on the activity of an additional output unit (or gating 'neuron'). The addition of this unit enabled agents to control whether or not linguistic inputs were recycled at the input layer.

The three conditions thus correspond to situations in which 1) agents 'hear' external instructions but do not repeat those instructions to themselves (no self-directed speech), 2) agents 'hear' external instructions and continuously repeat those instructions to themselves (continuous self-directed speech), and 3) agents 'hear' external instructions and autonomously decide whether or not to repeat those instructions back to themselves at each processing cycle (controlled self-directed speech).

Clowes and Morse found that in the control condition agents took longer to respond appropriately to any of the commands, and they additionally seemed unable to learn how to respond to different commands. In the other two conditions, agents were able to learn to respond appropriately to all of the commands, and they were additionally able to deal with multiple commands. Agents in the third condition (the one involving self-gated recycling of the linguistic inputs) were particularly successful at the task. These agents achieved the best overall performance in terms of evolutionary cost (i.e. the agents required the fewest number of generations in order to evolve successful performance profiles). The results thus hint at the importance of what Iizuka and Ikegami (2004) refer to as autonomous coupling – the ability of an agent to “spontaneously switch on or off its interaction with the environment” (pg. 283). It seems that when agents are allowed to exert some control over the extent to which both internal and external sources of information come to influence overt behaviour, then overall performance is enhanced (Clowes & Morse, 2005; Iizuka & Ikegami, 2004). This notion of autonomous coupling is thus closely related to the aforementioned notion of adaptive coupling. The difference is merely that autonomous coupling assumes that the coupling dynamics will be under the control of a specific agent, whereas the notion of adaptive coupling simply assumes that the coupling will be of cognitive benefit and makes no assumptions about the locus of control. In fact, in situations where there is selective pressure on agents to demonstrate good performance (as in the genetic algorithm study of Clowes and Morse), the timing of specific engagements with inner and outer sources of informational influence are likely to be driven towards optimal performance. In this case, autonomous coupling benefits overall performance and is synonymous with the notion of adaptive coupling.

We are now in a position to see why the temporal profile of adaptive coupling is unimportant. The studies of Clowes and Morse (2005) and Iizuka and Ikegami (2004) demonstrate that what is important is simply that the opportunity for adaptive coupling is available; the precise temporal details concerning the dynamics of the coupling process will largely depend on the nature of the task in which an agent is engaged (and may even depend on characteristics of the particular agent concerned). In the context of collective problem solving, it may sometimes be useful to have a period of independent (socially-decoupled) exploration before information exchange takes place. On other occasions it may be the *initial* exposure to other's ideas that sets our thought processes off on productive solo journeys to otherwise unreachable parts of some intellectual terrain. The main point, for present purposes, is simply that nothing restricts adaptive coupling to a particular point in time. Dynamic networks permit flexible forms of engagement and disengagement with a broad array of resources, and what is important for network-enabled cognition is merely the opportunity to modify network structures in ways that are adaptively aligned with the goals of both individual and collective cognition.

Future Research

Previous sections have highlighted a number of areas of research that are relevant to the notion of collective cognition. It should be clear, however, that many pressing research issues and questions remain. In this section, we present a number of areas for future research in the area of network-enabled collective cognition.

Cultural Models and Collective Cognition

Earlier in the paper we saw that coupled belief systems could be modelled by adopting Kauffman’s NK(C) model of co-evolutionary dynamics. The behaviour of these coupled belief systems in response to the manipulation of various network-level variables (e.g. network structure) constitutes an obvious target for future research, with a particularly interesting focus of attention being the inter-play between various sources of influence on the cognitive dynamics of multi-agent systems.

We saw that the cognitive dynamics of individual agents are governed by forces operating at a number of levels: the individual level, the social level and perhaps even the cultural level. This gives rise to a rich source of issues to be addressed by future research. The following are just some examples:

1) **How does the profile of inter-agent interaction at the social network level contribute to the emergence, modification and dissolution of cultural groups?** The aim here is to understand how the dynamics of inter-agent interaction (as perhaps determined by time-variant patterns of functional connectivity at the social network level) contribute to the evolution of distinct cultural models. Recall that we can use the methods developed as part of CNA to construct cultural models on the basis of information about individual agent beliefs. In the case of our simulation model, we have complete access to the beliefs of all agents in the simulation, and we can thus automatically construct and modify cultural models based on the changing profile of beliefs possessed by all the agents. We now encounter an interesting set of questions relating to the emergence of distinct cultural models in a multi-agent system. If, for example, we begin the simulation with multiple cultural models (reflecting the presence of culturally-significant groupings of agents), how will the cultural models change over time in response to the dynamics of information flow and influence at the social network level? Will all agents come to adopt a single model (perhaps reflecting a process of cultural convergence or cultural assimilation), or will the models merge to form a single hybrid cultural model (perhaps reflecting a process of cultural hybridization)? Perhaps the original cultural models will be preserved, but their boundaries at the social level will be highly porous (i.e. different agents at different times will switch between different cultural communities). Clearly, the combination of social networks, coupled belief systems and automated cultural model construction provides us with a rich source of simulation opportunities, with potential implications for our understanding of the evolution of cultural groups and the dynamics of culturally-entrenched belief systems[25].

2) **How do social networks affect the vulnerability of cultural groups to cognitive change?** The idea here is to examine how aspects of the social network (e.g. the network structure) affect the susceptibility of a cultural community (the agents associated with a particular cultural model) to cognitive change given some particular kind of external perturbation. Perhaps the most obvious perturbation here concerns exposure to agents from another cultural group. Thus, if a group of agents is exposed to agents associated with another cultural model, will they resist 'conversion' to the 'new way of thinking'? And how will this susceptibility to conversion relate to the structural features of the social network in which the respective groups of agents are embedded?

3) **How do cultural-level influences affect individual cognitive change?** The aim here is to explore the interaction between forces for cognitive change at the individual, social and cultural level. One possible simulation involves the creation of additional sources of

[25] There is an interesting set of questions here regarding the 'representation' of cultural groups within the social network. For example, one issue of interest concerns the 'spatial' distribution of agents from the same cultural group within a culturally-heterogeneous social network (i.e. a social network yielding more than one cultural model). Members of a particular cultural group may be spatially clustered within the social network, with high levels of connectivity between the group members, but we can also imagine situations where individuals from a cultural group are somewhat more 'dispersed' throughout the network.

influence between an individual's beliefs and the beliefs of the cultural model to which they are associated (the influence could be represented via the inclusion of additional epistatic linkages, perhaps with higher weightings than those associated with other linkages[26]). The cultural models could still be subject to automatic creation and modification (as discussed above); however, in this case, the cultural models would provide an additional source of influence on the cognitive dynamics of individual agents – agent cognition would be influenced by the aggregate beliefs of the cultural group to which they belonged.

4) **Do forces at the inter-cultural level contribute to the cognitive stability of culturally-heterogeneous groups?** The aim here is to examine the notion that forces operating at the cultural level (i.e. the level of cultural models) contribute to a form of 'ideological equilibrium' in which the cognitive outlooks of different cultural groups are stabilized by the emergence of Nash equilibria at the inter-cultural level.

Dynamic Networks and Collective Problem-Solving

The preliminary work on dynamic networks and collective problem-solving reported in this paper suggests that constructive networks could deliver performance benefits, relative to their more statically-configured counterparts. There are clearly a number of ways in which these results could be extended. One possible extension concerns the ways in which links are added to the network. Thus, aside from the development of randomly-generated networks in which links are added between randomly selected nodes, one could imagine 'growing' networks using a variety of other algorithms, for example, preferential selection processes in which nodes associated with either the most links or the fittest solutions are preferentially selected for link attachment.

Another extension to the work on dynamic networks and collective problem-solving concerns the role that 'information velocity' plays in determining collective performance. A variety of research results seem to suggest that highly efficient modes of inter-agent communication (in which information is rapidly disseminated to all parts of a social network) do not necessarily deliver the best long-term performance outcomes. Instead, more restricted modes of inter-agent communication may be beneficial, at least in some problem-solving contexts. Now, clearly there are a number of ways in which the efficiency of inter-agent communication could be affected. One is to structure the network in such a way that the rate of information flow and influence is effectively reduced; another is to introduce factors that limit the tendency of agents to process or pass on information that is supplied by their network neighbours. Recall, for example, the role that conditional probabilities played in influencing the extent of belief propagation through an agent network in Glinton et al's (2010) study. The conditional probabilities reflected the *credibility* that agents assigned to the information supplied by their peers, and this was sufficient to determine the extent to which an agent was influenced by its neighbours; i.e. whether it was inclined to revise its own belief to match that of its neighbours. The upshot is that credibility judgements contribute to an *effective* rate of information transmission that is largely independent of the topological characteristics of the network in which such transmission takes place.

[26] The basic idea here is to represent the cultural model as the belief system of another agent (a cultural model agent), to which all agents in the cultural group are connected. The addition of higher weightings between the epistatic ties of individual-level belief structures and cultural-level structures means that the cultural model exerts a disproportionate influence on the beliefs of cultural group members.

In human social networks a variety of psychosocial factors may influence the effective rate of information transmission. One such factor is, of course, the level of trust that exists between agents. When agents have high trust in one another, we may expect them to exert greater levels of influence than in cases of low trust, and this is likely to mean that information velocity is greater in high trust situations. If true, this would hint at a potentially interesting hypothesis concerning the adaptive value of *distrust* for collective problem-solving performances. Because distrust may effectively retard the rate of information dissemination through a network, collective problem-solving performances may be better in situations where agents initially *distrust* one another, or are at least somewhat circumspect about what others tell them. Perhaps a *dynamic* profile of trust evaluation could be imagined in which initially low levels of trust are supplanted by progressively greater levels of trust as agents begin to interact and communicate with each other. Such a profile would clearly deliver roughly the same kind of pattern of information flow and influence that we see in the case of the constructive network simulations reported earlier in this paper[27].

All of this serves to bring to light a specific hypothesis concerning the adaptive value of evolving trust relationships in collective problem-solving situations. The hypothesis is that trust serves an important role in determining the adaptive coupling of problem-solving agents into increasingly close-knit information processing ensembles. Initial levels of distrust, reflecting perhaps the initial caution people bring to new social situations, are not necessarily to be regarded as maladaptive when it comes to collective problem-solving. Sometimes distrust may play an adaptive role in configuring the functional connectivity of a network in a way that best meliorates collective cognitive processing. In some cases, as in the case of the working memory limitations of young infants (Kail, 1984) (recall Elman's (1993) comments in regard to the limited memory solution for neural networks), distrust may prevent a community of agents from prematurely converging on a sub-optimal solution (Lazer & Friedman, 2007) or forming an inaccurate shared interpretation of some external state-of-affairs (Hutchins, 1995). Rather than see distrust as something uniformly detrimental to collective cognition – something to be abhorred and ideally eliminated by future technological innovation – perhaps we should not be so quick to yield to our intuitions. For there may just be a brighter side to distrust. It may be that distrust enables groups to adaptively regulate the temporal profile of network-mediated information flow and influence in a way that facilitates the long-term realization of high quality collective cognitive outcomes. Perhaps every cloud does have a silver lining after all!

Collective Problem-Solving

The problem with silver linings is that they often come attached to rather ominous looking grey clouds. The clouds in the case of the foregoing discussion concern the generalizability of empirical results that are obtained in the rather restricted types of problem-solving domain used by Mason et al (2005) and Lazer and Friedman (2007). The fact is that both Mason et al (2005) and Lazer and Friedman (2007) use a specific type of task (involving collective search for optimal solutions) that is rather unlike that seen in many cases of real-world collective problem-solving. In order to improve the generalizability of the results, we need to look at the role of factors like network structure in

[27] This dynamic profile of inter-agent trust evaluation effectively sets up the conditions for the emergence of constructive networks (networks in which the conduits of information flow and influence gradually emerge across time). It should be clear that such networks are on a functional par with those described in the context of the dynamic network simulations reported earlier.

other types of problem-solving activity, particularly those involving some degree of task specialization and inter-agent dependence. Further research will be needed to look at the various types of problem-solving task in which social groups participate (perhaps developing a formal taxonomy of task types) and develop models of these tasks for use in future computer simulation studies.

Network Studies on Shared Interpretation

Recall Hutchins' (1995) work on the role played by inter-agent communication in enabling a community of agents to arrive at an accurate (and sometimes not so accurate) shared interpretation of some external state-of-affairs. Such work has potential implications for our understanding of how a network affects a group's ability to deal with poor quality information, for example, information that is ambiguous, inaccurate, inconsistent or incomplete. In thinking about possible extensions to Hutchins' work, it will no doubt be important to reflect on the relevance of Hutchins' studies to our notions of both shared situation awareness (Nofi, 2000) and shared understanding (Smart et al., 2009a; Smart et al., 2009b). It will also be important to assess whether factors like time-variant trust evaluations and dynamic network topologies deliver the same profile of performance benefits in the case of shared interpretations as they do in other cases of collective problem-solving. Finally, there are some potentially very interesting linkages between the work on shared interpretation and the work on coupled belief systems and cultural models. In particular, the notion of coupled belief systems seems to capture elements of the internal drive for consistent interpretations seen in the case of constraint satisfaction networks, as well as the influence exerted by inter-agent relationships. Furthermore, the notion of cultural models reminds us that within any significantly sized population we may encounter culturally-significant groupings that differ with regard to their beliefs, attitudes and values. Applying the notion of cultural models to Hutchins' simulations encourages us to think about groups of agents that have specific biases and predispositions with regard to their interpretation of environmental information. That is, we may see multiple groups of agents that are inclined to interpret common bodies of information in *different* ways based on their membership of specific cultural groups. By introducing the notion of cultural models into network simulations of shared interpretation, we thus potentially approximate the kind of socio-cognitive dynamics we see in situations involving the cooperative (and sometimes not so cooperative) interaction of culturally-disparate communities. This work is thus highly relevant to research programmes, such as the International Technology Alliance[28] (Preece & Sieck, 2007), which attempt to further our understanding of the forces and factors that influence collective cognitive processing in multi-cultural military coalitions.

Conclusion

This has been a long paper but an important one in terms of our understanding of network-enabled cognition. From our initial analysis of belief propagation in heterogeneous agent communities we have now examined a variety of ways in which the cognitive profile of multi-agent systems can be both enabled (and sometimes disabled) by the dynamics of network-mediated information flow and influence. Moving on from the work of Glinton et al (2010) and Parachuri et al (2010) we saw how a specific model of biological evolution, one which was initially developed to support our understanding of complex adaptive systems, could be used to potentially advance our understanding

[28] http://www.usukita.org/

of cognitive state fluctuations in social networks. In particular, we suggested that Kauffman's (1993, 1995) NK(C) model could be used to support our understanding of cognitive flux in epistatically-linked belief systems. The key idea here concerned how the genetic notion of epistasis could be applied to the problem of representing the inter-dependencies between specific beliefs at both the individual and collective levels. We also saw how the NK(C) model could be applied to belief systems at the cultural level, and we advanced the, admittedly speculative, proposal that the relative stability of particular ideologies (cultural-level belief systems) could be understood in terms of the emergence of Nash equilibria. Such ideas provide the basis for a variety of future computer simulation studies in which the focus of analysis is on the way in which networks contribute to cognitive dynamics at the individual, social and cultural levels.

Aside from the role of networks in cognitive state fluctuations, we also examined work exploring the effects of network-level variables on collective problem-solving performances. Here we encountered the idea that time-variant changes in network architecture permit flexible forms of adaptive coupling between the elements of a larger information processing system. Dynamic networks, we argued, enable agents to regulate their exposure to social information flow and influence in a way that benefits long-term collective performance.

The notion of adaptive coupling is of potential significance not just for our understanding of collective cognition; it also enables us to perhaps better understand the profile of much individual cognition. Thus just as the notion of adaptive coupling in social networks may help us better understand the roots of collective cognitive success, so too our understanding of individual cognition may be enhanced by recognizing the way in which material resources are flexibly factored into episodes of individual cognitive processing. The key idea here is that human cognitive processing is not necessarily something that relies solely on the dynamics of neural information processing; instead, cognition may sometimes be environmentally-extended, entangling all manner of representational and computational resources into complex time-variant nexuses of cognitively-relevant information processing. The result is what Wilson and Clark (2009) refer to as transient extended cognitive systems (TECS):

> *"A TECS is a soft-assembled whole that meshes the problem-solving contributions of the human brain and central nervous system with those of the (rest of the) body and various elements of the local cognitive scaffolding" (pg. 65)*

Applying these ideas to the realm of large-scale information networks and network-enabled devices gives us the notion of what we have referred to, in other work, as the 'network-extended mind' (Smart et al., 2010), a notion that is just as applicable to collective forms of cognition as it is to more individual, agent-centred forms of cognitive processing. In fact, rather than see collective cognition as something that is wholly distinct from individual cognition, perhaps we can now begin to see commonalities between the two phenomena. The notion of adaptive coupling, in conjunction with a network-based approach to modelling, simulation and analysis, enables us to see both individual and collective forms of cognition as *both* resulting from the flexible integration of a variety of resources into dynamically-configured networks. In fact, perhaps it no longer makes sense to talk of individual and collective forms of cognition; for all cognitive processing is, in some sense, collective. Rather than see our biological selves as the sole point source for all our individual cognitive successes and failures, perhaps we should learn to adopt a somewhat different standpoint – one that gives adequate emphasis and recognition to the role played by networks that transcend the traditional borders of skin and skull. And from such a standpoint, perhaps we can begin to see both individual and collective forms of cognition as on a functional par; for they both stem from the largely ephemeral webs of information flow and influence that are spun around the elements of our biological, social and technological worlds.

References

Adams, F., & Aizawa, K. (2001) The bounds of cognition. *Philosophical Psychology*, 14(1), 43-64.

Adams, F., & Aizawa, K. (2008) *The Bounds of Cognition*. Blackwell, Oxford, UK.

Ahouse, J. J., Bruderer, E., Gelover-Santiago, A., Konno, N., Lazer, D., & Veretnik, S. (1992) Reality kisses the neck of speculation: A report from the NKC workgroup. In L. Nadel & D. Stein (Eds.), *1991 Lectures in Complex Systems, SFI Studies in the Science of Complexity* (Vol. 4, pp. 331-353). Addison-Wesley, Reading, Massachusetts, USA.

Ash, T. (1989) Dynamic node creation in backpropagation networks. *Connection Science*, 1(365-375).

Bak, P., Tang, C., & Wiesenfeld, K. (1983) Self-organized criticality: an explanation of 1/f noise. *Physical Review Letters*, 59, 381-384.

Barabasi, A.-L. (2002) *Linked: The New Science of Networks*. Perseus Publishing, Cambridge, Massachusetts, USA.

Barnier, A. J., Sutton, J., Harris, C. B., & Wilson, R. A. (2008) A conceptual and empirical framework for the social distribution of cognition: The case of memory. *Cognitive Systems Research*, 9(1-2), 33-51.

Buchanan, M. (2002) *Nexus: Small Worlds and the Groundbreaking Science of Networks*. W. W. Norton & Company, New York, USA.

Clark, A. (2008) *Supersizing the Mind: Embodiment, Action, and Cognitive Extension*. Oxford University Press, New York, USA.

Clark, A. (2010) Coupling, Constitution and the Cognitive Kind: A Reply to Adam and Aizawa. In R. Menary (Ed.), *The Extended Mind*. MIT Press, Cambridge, Massachusetts, USA.

Clowes, R. W., & Morse, A. F. (2005) *Scaffolding cognition with words*. Fifth International Workshop on Epigenetic Robotics: Modeling Cognitive Development in Robotic Systems, Lund, Sweden.

Diehl, M., & Stroebe, W. (1987) Productivity loss in brainstorming groups: toward the solution of a riddle. *Journal of Personality and Social Psychology*, 53, 497-509.

Edelman, G. M. (1987) *Neural Darwinism: The Theory of Neuronal Group Selection*. Basic Books, New York, New York, USA.

Elman, J. (1990) Finding structure in time. *Cognitive Science*, 14, 179-211.

Elman, J. L. (1993) Learning and development in neural networks: The importance of starting small. *Cognition*, 48(1), 71-99.

Fahlman, S. E., & Lebiere, C. (1990) The cascade-correlation learning architecture. In D. S. Touretzky (Ed.), *Advances in Neural Information Processing Systems 2*. Morgan Kaufmann, San Mateo, California, USA.

Festinger, L. (1957) *A Theory of Cognitive Dissonance*. Stanford University Press, Stanford, USA.

Glinton, R., Paruchuri, P., Scerri, P., & Sycara, K. (2010) Self-Organized Criticality of Belief Propagation in Large Heterogeneous Teams. In M. J. Hirsch, P. M. Pardalos & R. Murphey (Eds.), *Dynamics of Information Systems: Theory and Applications*. Springer, Berlin, Germany.

Harnad, S. (1999) The future of scholarly skywriting. In A. Scammell (Ed.), *I in the Sky: Visions of the Information Future*. Aslib, London, UK.

Harnad, S., & Dror, I. E. (2006) Distributed cognition: Cognizing, autonomy and the Turing Test. *Pragmatics & Cognition*, 14(2), 209-213.

Hinsz, V., Tindale, R., & Vollrath, D. (1997) The emerging conceptualization of groups as information processors. *Psychological Bulletin*, 121, 43-64.

Hutchins, E. (1995) *Cognition in the Wild*. MIT Press, Cambridge, Massachusetts, USA.

Iizuka, H., & Ikegami, T. (2004) Simulating autonomous coupling in discrimination of light frequencies. *Connection Science*, 16(4), 283-299.

Janis, I. L. (1982) *Victims of Groupthink* (2nd ed.). Houghton Mifflin, Boston, USA.

Kail, R. (1984) *The Development of Memory*. W. H. Freeman, New York, New York, USA.

Kauffman, S. (1993) *The Origins of Order: Self-Organization and Selection in Evolution*. Oxford University Press, New York, USA.

Kauffman, S. (1995) *At Home in the Universe: The Search for Laws of Self-Organization and Complexity*. Penguin Books Ltd, London, England.

Lazer, D., & Friedman, A. (2006) *The dark side of the small world: how efficient information diffusion drives out diversity and lowers collective problem solving ability.* Kennedy School of Government, Harvard University, Cambridge, MA, USA.

Lazer, D., & Friedman, A. (2007) The Network Structure of Exploration and Exploitation. *Administrative Science Quarterly*, 52(4), 667-694.

Marder, E., & Bucher, D. (2007) Understanding circuit dynamics using the stomatogastric nervous system of lobsters and crabs. *Annual Review of Physiology*, 69, 291-316.

Mason, W. A., Jones, A., & Goldstone, R. L. (2005) *Propagation of innovations in networked groups*. 27th Annual Conference of the Cognitive Science Society, Stresa, Italy.

Monge, P. R., & Contractor, N. S. (2003) *Theories of Communication Networks*. Oxford University Press, New York, USA.

Mozer, M., & Smolensky, P. (1989a) Skeletonization: A technique for trimming the fat from a network via relevance assessment. In D. Touretzky (Ed.), *Advances in Neural Information Processing Systems I* (pp. 107-115). Morgan Kaufmann, San Mateo, CA, USA.

Mozer, M. C., & Smolensky, P. (1989b) Using Relevance to Reduce Network Size Automatically. *Connection Science*, 1(2), 3-16.

Nofi, A. A. (2000) *Defining and Measuring Shared Situational Awareness.* Center for Naval Analyses, Alexandria, Virginia, USA. (Ref: CRM D0002895.A1/Final)

Parachuri, P., Glinton, R., Sycara, K., & Scerri, P. (2010) Effect of humans on belief propagation in large heterogeneous teams. In M. J. Hirsch, P. M. Pardalos & R. Murphey (Eds.), *Dynamics of Information Systems: Theory and Applications*. Springer, Berlin, Germany.

Preece, A., & Sieck, W. R. (2007) The International Technology Alliance in Network and Information Sciences. *IEEE Intelligent Systems*, 22(5), 18-19.

Quartz, S. (1999) The constructivist brain. *Trends in Cognitive Sciences*, 3(2), 48-57.

Quartz, S., & Sejnowski, T. (1997) The neural basis of cognitive development: A constructivist manifesto. *Behavioral and Brain Sciences*, 20(4), 537-556.

Rasmussen, L. J., Sieck, W. R., & Smart, P. R. (2009) What is a Good Plan? Cultural Variations in Expert Planners' Concepts of Plan Quality. *Journal of Cognitive Engineering and Decision Making*, 3(3), 228-253.

Rowlands, M. (2006) *Body Language: Representation in Action*. MIT Press, Cambridge, Massachusetts, USA.

Schultz, T. R., & Schmidt, W. C. (1991) *A cascade-correlation model of balance scale phenomena*. 13th Annual Conference of the Cognitive Science Society, Chicago, Illinois, USA.

Selverston, A. (1995) Modulation of circuits underlying rhythmic behaviors. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, 176(2), 139-147.

Sieck, W. R., & Rasmussen, L. (2007) *Cultural Network Analysis*. 1st Annual Conference of the International Technology Alliance (ACITA), Maryland, USA.

Sieck, W. R., Rasmussen, L., & Smart, P. R. (2010) Cultural Network Analysis: A Cognitive Approach to Cultural Modeling. In D. Verma (Ed.), *Network Science for Military Coalition Operations: Information Extraction and Interaction*. IGI Global, Hershey, Pennsylvania, USA.

Simpkins, B., Sieck, W. R., Smart, P. R., & Mueller, S. T. (2010) *Idea Propagation in Social Networks: The Role of 'Cognitive Advantage'*. 1st ITA Workshop on Network-Enabled Cognition: The Contribution of Social and Technological Networks to Human Cognition, Maryland, USA.

Smart, P. R., Engelbrecht, P. C., Braines, D., Strub, M., & Giammanco, C. (2010) The Network-Extended Mind. In D. Verma (Ed.), *Network Science for Military Coalition Operations: Information Extraction and Interaction*. IGI Global, Hershey, Pennsylvania, USA.

Smart, P. R., Huynh, T. D., Mott, D., Sycara, K., Braines, D., Strub, M., Sieck, W. R., & Shadbolt, N. R. (2009a) *Towards an Understanding of Shared Understanding in Military Coalition Contexts*. 3rd Annual Conference of the International Technology Alliance (ACITA'09), Maryland, USA.

Smart, P. R., Mott, D., Sycara, K., Braines, D., Strub, M., & Shadbolt, N. R. (2009b) *Shared Understanding within Military Coalitions: A Definition and Review of Research Challenges*. Knowledge Systems for Coalition Operations (KSCO'09), Southampton, UK.

Stasser, G., & Titus, W. (1985) Pooling of unshared information in group decision making: biased information sampling during group discussion. *Journal of Personality and Social Psychology*, 48, 1467-1478.

Watts, D. J. (2002) A simple model of global cascades on random networks. *Proceedings of the National Academy of Sciences of the United States of America*, 99(9), 5766.

Watts, D. J. (2003) *Six Degrees: The Science of a Connected Age*. William Heinemann, London, UK.

Wilson, R. A., & Clark, A. (2009) Situated Cognition: Letting Nature Take its Course. In M. Ayded & P. Robbins (Eds.), *Cambridge Handbook of Situated Cognition*. Cambridge University Press, Cambridge, UK.

Zsambok, C. E., & Klein, G. A. (Eds.) (1997) *Naturalistic Decision Making*. Lawrence Erlbaum Associates, Mahwah, New Jersey.