Heterogeneity and Robustness in Social Learning

Jonathan Lawry

Department of Engineering Mathematics,
University of Bristol,
Bristol, BS8 1TW, UK
Email: j.lawry@bristol.ac.uk

Abstract

Social learning is an important collective behaviour in many biological and artificial systems. We investigate a model of social learning which combines two distinct processes, one relating to how individuals adapt their beliefs as a result of interacting with their peers, and one relating to when they search for and how they learn directly from evidence. For each process we introduce conservative and open-minded behaviours and combine these to obtain four social learning behaviour types. A simple truth-seeking task is considered and a three-valued model of belief states is adopted. By means of difference equation models and agent-based simulations we then investigate the performance of the different learning behaviours. We show that certain heterogeneous mixtures of behaviours result in the most robust performance for a variety of learning rates and initial conditions, and that such mixtures are well suited for social learning in dynamic environments.

Introduction & Background

Social learning is fundamental to collective decision making in many biological systems and on social networks, as well as being crucial for applications in swarm robotics and multi-agent AI. In contrast to individual learning in which agents operate alone when collecting evidence and making inferences, in social learning individuals are also influenced by observation of, or interaction with, others in a socially connected population (Heyes, 1994). This means that population level consensus or polarisation is ultimately an emergent effect of individual learning that depends on learning behaviour as well as the nature of the problem, the level and type of connectivity between agents and the amount and quality of evidence available.

In applications such as swarm robotics it is common to consider homogeneous populations of simple robots, all co-operating to reach a common goal or to solve a shared problem (Brambilla, 2013). However, in natural systems some form of heterogeneity is common within social groups, resulting in different behaviours and divisions of labour. In this paper we focus on the effect of heterogeneous learning behaviours on the efficacy of social learning. More specifically, we will consider different approaches to learning which vary in terms of how conservative or open-minded they are in their treatment of evidence and also in they way that individuals learn from their peers. By defining difference equation models and through multi-agent simulations we will demonstrate that a heterogeneous mixture of learning behaviours can enhance both accuracy and robustness in a simple type of social learning task.

The effect of heterogeneity on emergent behaviour in complex systems has been studied across a number of disciplines. For example, dating back to (Smith, 1776) heterogeneity in the division of labour has been seen as a key feature of market economies. Similar divisions of labour have also been observed and modelled in biology, particularly in social insects (Buckingham, 1911; O'Shea-Wheller et al, 2020). In this context threshold-response models have been used to explain how a certain form of individual heterogeneity enables social insect societies to regulate their environmental conditions (Theraulaz et al, 1998). These models assume that each individual responds to a given stimulus when its stimulus intensity exceeds their threshold, and that thresholds vary between individuals in the population. For instance, threshold-response models have been applied to collective adaptive ventilation behaviour in European honey bees, where individual honey bees exhibit fanning behaviour when the local temperature exceeds their individual threshold. Heterogeneous thresholds then enable a graded overall response by the hive to external temperature changes (Peters et al, 2019). Other types of behavioural heterogeneity are also present in social insect species. This includes varying levels of boldness in social spiders such as *Stegodyphus Dumicol* (Hunt et al, 2019), which impacts on overall attack speed.

Emergent behaviour resulting in consensus or polarisation on social networks of interacting heterogeneous agents has been investigated using opinion dynamics models (Baronchelli, 2018). For example, thresholds can be used as part of the continuous belief Hegselmann-Krause model to capture different behaviour types where agents are more or less open-minded. More specifically, agents are influenced by those neighbours on their social network with beliefs sufficiently similar to their own. Here similarity is quantified as a numerical measure between beliefs and two beliefs are deemed to be sufficiently similar if the degree of similarity between them exceeds a given threshold. In heterogeneous models, each agent makes these judgements based on their own threshold, and agents with lower thresholds can be considered as being more open minded. Opinion dynamics models aim to capture how agents on a social network influence each other's opinion over time, and hence they do not typically take account of the effect of evidence received directly from the environment. In contrast, in social learning individual agents must take account of both direct evidence and the beliefs of the other agents with whom they interact. This form of learning is important for collective decision making in biological systems such as social insects (Valentini et al, 2017), in swarm robotics and arguably in human societies e.g. as part of the scientific method (Douven & Kelp, 2011).

There is no reason to assume that for social learning homogeneous populations consisting of only one type of learning behaviour will always be optimal. For example, (Yaman et al., 2022) present simulation experiments for a simple multi-armed bandit problem indicating that best performance is achieved with a heterogeneous population consisting of independent (non-social) learners and two types of social learners, one type favouring learning from the most successful agents and the other tending to adopt the majority opinion. In this paper we argue that a heterogeneous mixture of behaviours can improve the robustness of social learning to varying initial conditions, and different rates of evidence and interactions between agents. It also enables social learning to be more adaptive in dynamic environments where the underlying state-of-the-world changes during the learning process (Prasetyo et al, 2019). We formulate learning in terms of two distinct but interacting processes; **evidential updating** according to which agents update their current beliefs based on direct evidence and **fusion** by which agents combine their beliefs with others with whom they are interacting. We describe two types of behaviour for each process resulting in four learning behaviours which agents can adopt in a social learning context.

Three-Valued Social Learning

We consider a simple social learning problem in which a population of agents attempt to determine the truth or falsity of a proposition relating to the environment in which they are operating. For instance, this could refer to the presence or absence of resources, e.g. food or shelter, in a certain location, or whether or not a particular option or choice is the best. Alternatively, in the context of swarm robotics applications such as search and rescue it could refer to the positions of casualties in a specified search area. A three-valued model of belief states is adopted where, at any time, an agent holds one of three beliefs; $\mathbf{f}$ indicating that they believe the proposition to be false, $\mathbf{u}$ indicating that they are uncertain or uncommitted, or $\mathbf{t}$ indicating that they believe the proposition is true. This approach has been shown to be an effective framework in social learning, which is robust to environmental noise and scalable to multiple propositions (Crosscombe & Lawry, 2017). Furthermore, for the best-of-n problem in swarm robotics the inclusion of the intermediate uncertain belief state $\mathbf{u}$ has been shown to improve robustness to the presence of malfunctioning robots in the swarm (Crosscombe et al, 2017). In addition, there is evidence that for some biological systems the presence of uncommitted individuals can help to facilitate population level consensus in decision making (Couzin et al, 2011).

In the following we propose different fusion and updating behaviours in the three-valued setting and formulate difference equation models for a totally collected well-mixed population in each case (Parker & Zhang, 2009). We consider time-stepped models where in each time-step agents attempt to undertake fusion followed by evidential updating. Difference equation models can then be formulated in terms of the proportion of agents in a large population holding each belief. More formally, we let:

$$\mathbf{P}_t = (P_t(\mathbf{f}), P_t(\mathbf{u}), P_t(\mathbf{t}))$$

denote the proportions of the three belief states in the agent population at time t. Macro-level learning can then be modelled as a difference equation of the form:

$$\mathbf{P}_{t+1}^T = U(F^{\mathbf{P}_t}\mathbf{P}_t^T) = (UF^{\mathbf{P}_t})\mathbf{P}_t^T$$

where U and $F^{\mathbf{P}}$ are matrices of transition probabilities for the updating and fusion processes respectively. These are 3×3 matrices of the form:

$$\begin{pmatrix} P(\mathbf{f}|\mathbf{f}) & P(\mathbf{f}|\mathbf{u}) & P(\mathbf{f}|\mathbf{t}) \\ P(\mathbf{u}|\mathbf{f}) & P(\mathbf{u}|\mathbf{u}) & P(\mathbf{u}|\mathbf{t}) \\ P(\mathbf{t}|\mathbf{f}) & P(\mathbf{t}|\mathbf{u}) & P(\mathbf{t}|\mathbf{t}) \end{pmatrix}$$

where, for example, $P(\mathbf{u}|\mathbf{f})$ denotes the probability that an agent currently in belief state $\mathbf{f}$ will transition to $\mathbf{u}$. In general, the transition probabilities for fusion are dependent on the current proportions of belief states in the population since they need to take account of how likely it is that a given agent will interact with another agent holding any of the different belief states. We denote this by the superscript in $F^{\mathbf{P}}$. On the other hand, for evidential updating, transition probabilities are independent of current belief state proportions.

Behaviour Types

Fusion Behaviours

We propose a simple three-valued model of a process whereby agents look to interact with other agents so as to be informed of and learn from their opinions. Specifically, within a time step each agent, acting as a receiving agent, attempts to interact with one other randomly selected agent,

	f	**u**	**t**
f	**f**	**f**	**u**
u	**f**	**u**	**t**
t	**u**	**t**	**t**

Table 1: Table for the adventurous operator. Each cell is the updated belief for a given current belief (row) and received belief (column)

	f	**u**	**t**
f	**f**	**u**	**u**
u	**u**	**u**	**u**
t	**u**	**u**	**t**

Table 2: Table for the cautious operator. Each cell is the updated belief for a given current belief (row) and received belief (column)

acting as a transmitting agent. Such an interaction occurs with probability σ (the fusion rate), quantifying limitations on communication within the population. All agents are able to adopt both receiving and transmitting roles. If no interaction takes place then the agent maintains their current belief as a default. As a result of such an interaction the receiving agent adapts their current belief state based on the belief state of the transmitting agent, according to a set of simple rules which can be represented in the form of a truth table. Here we propose two types of fusion behaviour, one more adventurous where agents are willing to adopt strong opinions expressed by their peers, and one more cautious in which agents are led to doubt their own opinions when others express uncertainty.

In **adventurous fusion** *certainty dominates over uncertainty* so that if the receiving agent with belief state $\mathbf{u}$ interacts with a transmitting agent with committed belief states $\mathbf{t}$ or $\mathbf{f}$, they adopt the latter according to table 1. Otherwise, inconsistency between the transmitting and receiving agents' truth values, i.e. $\mathbf{t}$ vs $\mathbf{f}$ or vice versa, results in the receiving agent changing their belief to uncertain. In this way strong disagreement between the transmitting and receiving agents results in the latter doubting their current belief. Consequently, we have the following transition probabilities for adventurous fusion:

$$F_A^{\mathbf{P}} = \begin{pmatrix} 1-\sigma P(\mathbf{t}) & \sigma P(\mathbf{f}) & 0 \\ \sigma P(\mathbf{t}) & 1-\sigma(P(\mathbf{t})+P(\mathbf{f})) & \sigma P(\mathbf{f}) \\ 0 & \sigma P(\mathbf{t}) & 1-\sigma P(\mathbf{f}) \end{pmatrix}$$

For illustrative purposes we derive one of these transition probabilities as follows: An agent currently in belief state $\mathbf{t}$ can remain in that state after adventurous fusion in one of two ways. They can fail to interact with any other agent and hence maintain their current belief as a default. This will occur with probability $1-\sigma$. Alternatively, they can succeed in interacting with another agent who either also has belief $\mathbf{t}$ or has belief $\mathbf{u}$ (see table 1). This has probability $\sigma(P(\mathbf{t})+P(\mathbf{u})) = \sigma(1-P(\mathbf{f}))$. Hence,

$$P(\mathbf{t}|\mathbf{t}) = 1-\sigma+\sigma(1-P(\mathbf{f})) = 1-\sigma P(\mathbf{f})$$

In **cautious fusion** *uncertainty dominates over certainty* so that if a receiving agent with committed belief state $\mathbf{t}$ or $\mathbf{f}$ interacts with a transmitting agent with uncertain truth state $\mathbf{u}$ they will abandon their committed position and change their belief to $\mathbf{u}$ according to table 2. Consequently, we have the following transition probabilities for cautious fusion:

$$F_C^{\mathbf{P}} = \begin{pmatrix} 1-\sigma+\sigma P(\mathbf{f}) & 0 & 0 \\ \sigma(1-P(\mathbf{f})) & 1 & \sigma(1-P(\mathbf{t})) \\ 0 & 0 & 1-\sigma+\sigma P(\mathbf{t}) \end{pmatrix}$$

Evidential Updating Behaviours

In this model, evidence corresponds to a statement of the truth value of the relevant proposition, either $\mathbf{t}$ or $\mathbf{f}$, resulting from a (hypothetical) investigation of the environment by the agent concerned, and influenced by the noise parameter ϵ. Without loss of generality, we assume here and throughout unless stated otherwise, that the correct truth value for the proposition is $\mathbf{t}$ and evidence will report this with probability $1-\epsilon$ while the erroneous truth value $\mathbf{f}$ will be reported with probability ϵ. On receiving evidence of this form an agent will update their current belief depending on one of two behaviour types. This updating only takes place if they find evidence once they decide to look, and this occurs with probability ρ (the evidence rate). If no evidence is found then the agent maintains their current belief as a default.

In **confident updating**, once agents are committed to belief state $\mathbf{t}$ or $\mathbf{f}$ they are sufficiently confident in their opinion that they cease looking for evidence. Therefore, since agents only look for evidence if they are in the uncertain state $\mathbf{u}$, evidential updating simply involves them changing their belief state to whichever is asserted by the evidence. This means that transition probabilities for the evidential updating process are constrained such that $P(\mathbf{t}|\mathbf{t}) = P(\mathbf{f}|\mathbf{f}) = 1$. See the flow diagram in figure 1a. The full matrix of transition probabilities for confident updating is given by:

$$U_C = \begin{pmatrix} 1 & \rho\epsilon & 0 \\ 0 & 1-\rho & 0 \\ 0 & \rho(1-\epsilon) & 1 \end{pmatrix}$$

According to **inquisitive updating** agents are always sufficiently curious to look for evidence no matter what is their current belief state. This means that evidence may be inconsistent with an agent's current belief i.e. $\mathbf{f}$ given $\mathbf{t}$ or vice versa. In this case the agent reverts to truth state $\mathbf{u}$. Otherwise, the agent simply adopts the truth state asserted by the evidence. See the flow diagram in figure 1b. This results in the following matrix of transition probabilities.

$$U_I = \begin{pmatrix} 1-\rho+\rho\epsilon & \rho\epsilon & 0 \\ \rho(1-\epsilon) & 1-\rho & \rho\epsilon \\ 0 & \rho(1-\epsilon) & 1-\rho\epsilon \end{pmatrix}$$

Again for illustration we derive one of the above transition probabilities. There are two ways that an agent currently in belief state $\mathbf{f}$ can remain in that state after inquisitive updating. They can fail to find evidence and this has probability $1 - \rho$. Alternatively, since we are assuming that the proposition is actually $\mathbf{t}$, they can find erroneous evidence. This has probability $\rho\epsilon$. Hence,

$$P(\mathbf{f}|\mathbf{f}) = 1 - \rho + \rho\epsilon$$

From a certain perspective it is possible to view inquisitive and confident updating as special kinds of explore and exploit strategies respectively. Confident updating preserves the committed belief states of $\mathbf{t}$ and $\mathbf{f}$, and this can be exploited during fusion (especially adventurous fusion) to drive consensus across the population. On the other hand, inquisitive updating encourages exploration in the form of a constant search for evidence by all agents. In general, identifying a good trade-off between explore and exploit behaviours has been found to be beneficial in a variety of learning contexts including reinforcement learning (Schäfer et al, 2021), multi-armed bandits (Slivkins, 2019), and evolutionary algorithms (Črepinšek et al, 2013).

If we now consider social learning where agents learn both through fusion and evidential updating, then taking the conjunction of the behaviour types described above for both of these processes, naturally results in four behaviour types for the combined learning process; **adventurous & confident (AC)**, **cautious & confident (CC)**, **adventurous & inquisitive (AI)** and **cautious & inquisitive (CI)**. In the following section we investigate the learning efficacy of homogeneous populations comprising solely of agents of each of these four behaviour types.

Properties of Behavioural Types

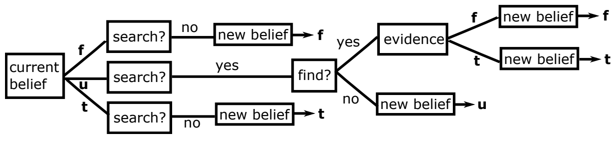
In this section we investigate the learning performance of the four behaviour types introduced above in a *homogeneous* setting, under significant noise $\epsilon = 0.3$, for varying fusion (σ) and evidence rates (ρ), and for different initial proportions of beliefs in the population. Figures 3 and 4 show heat maps of $P(\mathbf{t})$, i.e. the proportion of agents who have learnt the correct truth value, at $t = 1500$; evidence rates ρ and fusion rates σ vary in steps of 0.01 across the interval $(0, 1)$. Figure 3 is for a population of agents with initial beliefs such that 90% of agents begin the learning process committed to the incorrect belief (i.e., $\mathbf{f}$) and 10% to the correct belief i.e. $\mathbf{P}_0 = (0.9, 0, 0.1)$. This is clearly a challenging configuration of agent beliefs from which to initialise learning but we will argue that the capacity to learn effectively starting from this type of initial condition can be important when faced with dynamic environments where the underlying true state-of-the-world can change suddenly. Figure 4 assumes that all agents begin the learning process in the uncertain state $\mathbf{u}$ i.e. $\mathbf{P}_0 = (0, 1, 0)$. This scenario is particularly relevant for some applications of collective learning in robotics and autonomous systems when initial settings can be directly controlled.

For initial proportions $\mathbf{P}_0 = (0, 1, 0)$ it is clear from figure 4a that the best performance is obtained from a population consisting entirely of **AC** agents. Unsurprisingly, performance is better for initialisation $(0, 1, 0)$ than for $(0.9, 0, 0.1)$ for all four behaviour types, although the difference is small for cautious fusion agents (i.e. **CC** and **CI**). In general with the exception of **AI** initialised at $(0, 1, 0)$, all agents perform best when the evidence rate is relatively high compared to the fusion rate. Furthermore, although there are differences between the behaviour types regarding which combinations of σ and ρ give good performance, there are some interesting similarities, especially between **AC** and **CC**. In particular, figures 3a, 3b and 4b all show clear regions of good vs poor performance bounded by what seems to be the same functional relationship between ρ and σ. We can gain more insight into this by considering the stability of certain equilibrium points of the difference equation for these two behaviour types.

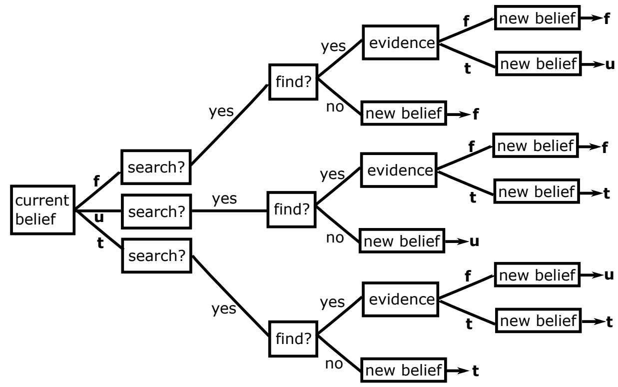
Note that both $(1, 0, 0)$ and $(0, 0, 1)$ are equilibrium (fixed) points of the **AC** and **CC** difference equations. Since these represent the situations in which the whole population of agents reach consensus about the incorrect and correct state-of-the-world respectively, it is therefore insightful to consider the stability of these equilibrium points under different noise conditions, fusion and evidence rates. This can be done by determining the Jacobian matrix for both difference equation models, evaluating it at the two fixed points, and considering the absolute values of the respective eigenvalues. If these values are all strictly less than 1 then the equilibrium point is stable, while it is unstable if any are strictly greater than 1. Accordingly, we find that for the **AC** and **CC** behaviour types there is a natural boundary dividing the (σ, ρ) parameter space into different stability regions given by:

$$\rho = \frac{\sigma}{1 + \sigma - 2\epsilon} = f(\sigma; \epsilon)$$

This then corresponds to the boundary for the regions observed in figures 3a, 3b and 4b. However, the exact nature of this division is different for **AC** and **CC**. For **AC**, $(0, 0, 1)$ is stable for all $\sigma, \rho \in (0, 1)$ and $\epsilon \in [0, 0.5)$, but $(1, 0, 0)$ is stable if $\rho < f(\sigma; \epsilon)$ and unstable if $\rho > f(\sigma; \epsilon)$ (see figure 2a). In contrast, for **CC** $(1, 0, 0)$ is unstable for all $\sigma, \rho \in (0, 1)$ and $\epsilon \in [0, 0.5)$, but $(0, 0, 1)$ is unstable if $\rho < f(\sigma; \epsilon)$ and stable if $\rho > f(\sigma; \epsilon)$ (see figure 2b). Hence, for both behaviour types with combinations of fusion and evidence rates such that $\rho > f(\sigma; \epsilon)$ we have that the incorrect consensus is an unstable equilibrium while the correct consensus is a stable equilibrium, and this helps to explain the good performance for both **AC** and **CC** in this region of the heat maps shown in figures 3a, 3b, 4a and 4b. For

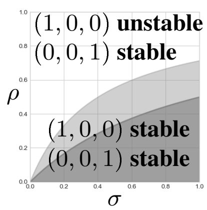


(a) Confident Updating

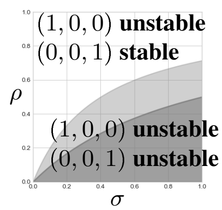


(b) Inquisitive Updating

Figure 1: Flow diagram showing the evidential updating process for the confident and inquisitive behaviours.



(a) **AC** behaviour type



(b) **CC** behaviour type

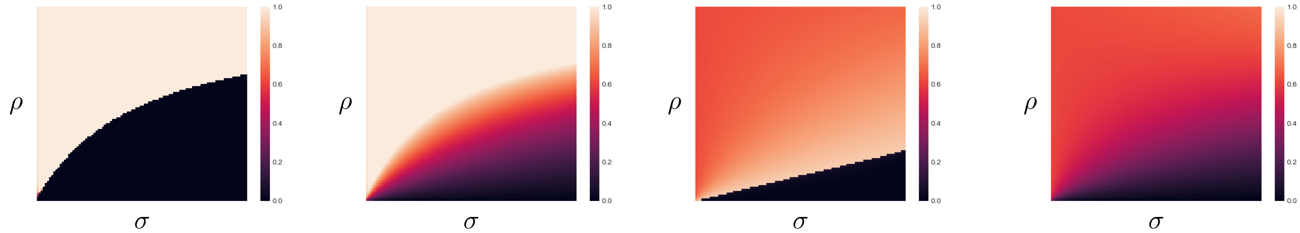
Figure 2: Stability of equilibrium points $(1, 0, 0)$ and $(0, 0, 1)$ with different combinations of fusion rate (σ) and evidence rates (ρ), for adventurous & confident and cautious & confident learning behaviour types. For both types there are two stability regions partitioned by $\rho = \frac{\sigma}{1+\sigma-2\epsilon}$, shown here for the noise values of $\epsilon = 0.3$ (light grey) and $\epsilon = 0$ (dark grey).

CC with combinations of fusion and evidence rates such that $\rho < f(\sigma; \epsilon)$ the instability of both equilibria seems to result in more mixed performance in this region of parameter space, especially for (σ, ρ) close to the boundary (see figures 3b,4b). In contrast, the stability of both equilibria in this region for **AC**, tends to result in convergence to $(1, 0, 0)$ (figure 3a) for initial condition close to that equilibrium point and to $(0, 0, 1)$ otherwise (figure 4a). Furthermore, notice that $f(0; \epsilon) = 0$ for all ϵ, and $f(\bullet; \epsilon) \leqslant f(\bullet, \epsilon')$ for $0 \leqslant \epsilon' \leqslant \epsilon < 0.5$. This monotonicity property implies that the area of the upper region of σ, ρ parameter space in which only the correct consensus is a stable equilibrium and which therefore tends to be associated with good performance, decreases as the noise increases, while the area of the lower region associated with mixed or poor performance increases. For example, contrast the light and dark grey regions in figure 2 corresponding to $\epsilon = 0.3$ and $\epsilon = 0$ respectively.

Space of Learning Behaviours

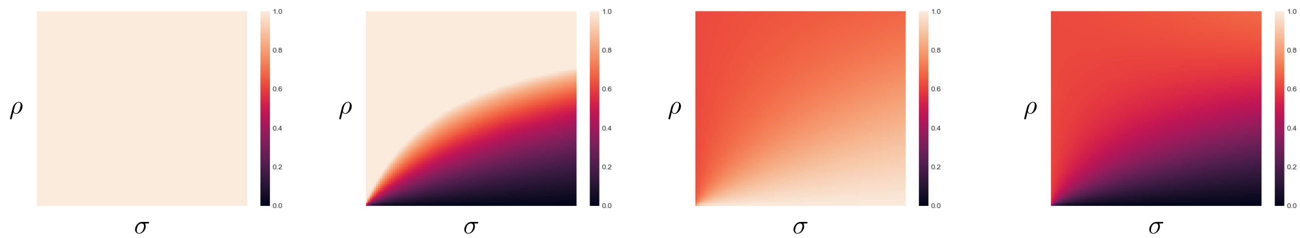
In the previous section we considered the collective learning performance of homogeneous populations consisting of each of the four main behaviour types. In this section we consider heterogeneity of behaviour types within a single population of agents. Initially, we fix the updating behaviour to be confident and consider a mixed population of **AC** agents, proportion w, and **CC** agents, proportion $1 - w$, for $w \in [0, 1]$. Figure 5 shows $P(\mathbf{t})$ at $t = 1500$ for initial condition $(0.9, 0, 0.1)$ for varying w and for different fusion and evidence rates. For the cases shown a $50/50$ mixture (i.e. $w = 0.5$) of **AC** and **CC** behaviours results in optimal performance. Indeed, the heat maps of $P(\mathbf{t})$ across σ and ρ shown in figure 6 indicate that this mixture of behaviour types performs consistently well for all combinations of fusion and evidence rate and for both initial proportions $(0.9, 0, 0.1)$ and $(0, 1, 0)$. The stability of the equilibrium points $(1, 0, 0)$ and $(0, 0, 1)$ also shed some light on why heterogeneous behaviour of this form is so effective, since for $w = 0.5$, $(0, 0, 1)$ is stable and $(1, 0, 0)$ is unstable for all $\sigma, \rho \in (0, 1)$ and $\epsilon \in [0, 0.5)$. However, for any mixture of only **AC** and **CC** behaviours the incorrect state-of-the-world $(1, 0, 0)$ is an equilibrium point and hence learning cannot take place for a population which begin with this incorrect consensus. In such cases, we hypothesise that including some inquisitive learners in the population may be helpful.

We now suppose that a proportion λ of the agent population are inquisitive regarding evidence while the remainder are confidence. Assuming that the behaviours for fusion and for evidential updating are allocated independently then this results in a mixture of the four behaviour types with proportions as given in table 3. Note that the allocation of behaviour types to agents can either be made at the beginning of the simulation and then fixed, or reallocated at every time step. For the latter we can think of agents independently



(a) **AC** behaviour type (b) **CC** behaviour type (c) **AI** behaviour type (d) **CI** behaviour type

Figure 3: Heat maps showing the proportion of correct beliefs with different combinations of fusion rate (σ) and evidence rates (ρ), for the four behaviour types initialised at $(0.9, 0, 0.1)$ i.e. with 90% of agents initially believing the wrong answer. Results are at $t = 1500$ and $\epsilon = 0.3$.



(a) **AC** behaviour type (b) **CC** behaviour type (c) **AI** behaviour type (d) **CI** behaviour type

Figure 4: Heat maps showing the proportion of correct beliefs with different combinations of fusion rate (σ) and evidence rates (ρ), for the four behaviour types initialised at $(0, 1, 0)$ i.e. with all agents initially uncertain. Results are at $t = 1500$ and $\epsilon = 0.3$.

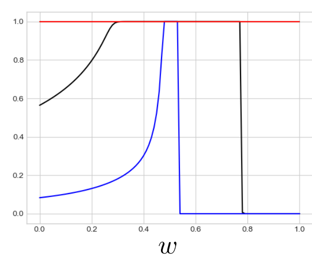
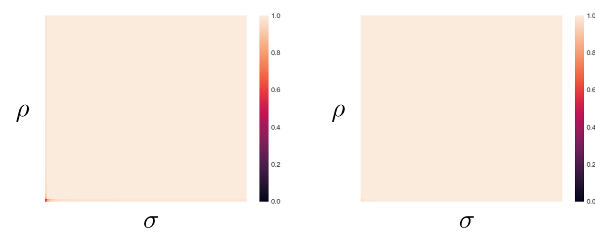


Figure 5: Proportion of correct beliefs for different proportions of **AC** (w) and **CC** ($1 - w$) behaviour types. Black line for $\sigma = 0.1$, $\rho = 0.1$, blue line for $\sigma = 0.9$, $\rho = 0.1$, green line for $\sigma = 0.1$, $\rho = 0.9$, and red line for $\sigma = 0.9$, $\rho = 0.9$. Results are for $\epsilon = 0.3$ at $t = 1500$ and for initial proportions $(0.9, 0, 0.1)$.



(a) Initialisation $(0.9, 0, 0.1)$. (b) Initialisation $(0, 1, 0)$.

Figure 6: Heat maps showing the proportion of correct beliefs with different combinations of fusion rate (σ) and evidence rates (ρ), for a $50/50$ mixture of **AC** and **CC**. Results are at $t = 1500$ and $\epsilon = 0.3$.

choosing a behaviour type at random according to the probabilities given in table 3 at every time step. Since we are assuming full mixing of totally connected agents the difference equation model does not distinguish between the two variations, but it will make a different to the implementation of the agent-based simulation in the following section.

Investigating performance across the space of parameter values for $(1, 0, 0)$ (figure 7) can now provide insight into which heterogeneous populations are robust to learning under the most challenging initial conditions. Figure 8 shows heat maps of $P(\mathbf{t})$ at $t = 1500$ for $\epsilon = 0.3$ and initial proportion $(1, 0, 0)$ for varying λ and w, and for different combinations of fusion and evidence rates. Taken together these suggest that good performance can be obtained for a low but non-zero proportion of inquisitive agents, e.g. $\lambda = 0.01$, and a $50/50$ split between adventurous and cautious agents, i.e. $w = 0.5$. In the next section, we show that this combination of agent behaviours is able to adapt to learning in a dynamic

	Adventurous w	Cautious $1-w$
Confident $1-\lambda$	$(1-\lambda)w$ **AC**	$(1-\lambda)(1-w)$ **CC**
Inquisitive λ	λw **AI**	$\lambda(1-w)$ **CI**

Table 3: Proportions of behaviour types in the population

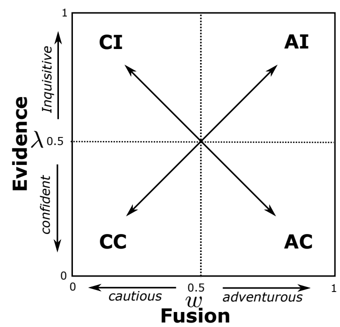


Figure 7: Space of parameter values (w, λ) representing different proportions of learning behaviour types.

environment in which the true state-of-the-world suddenly changes.

Agent-Based Model of Dynamic Environments

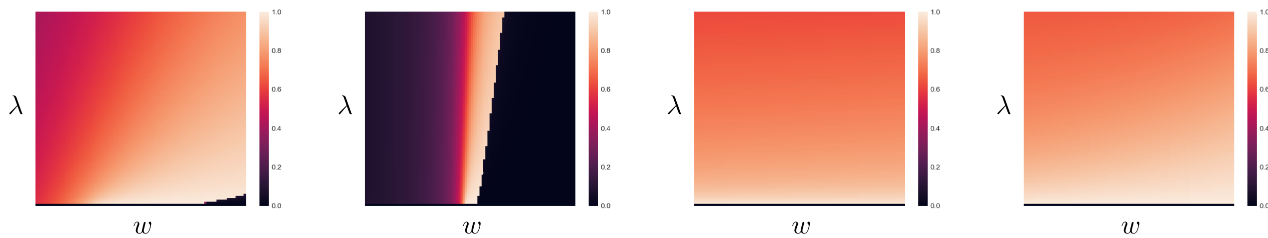
It is common for social learning to take place over a time-scale during which the state-of-world may change (Prasetyo et al, 2019). Here we consider a simple scenario in which the true state-of-the-world changes suddenly from **f** to **t** at time step $t = 500$. This scenario is investigated using an agent-based simulation rather than difference equation model of the type discussed above. A population of 200 agents are all initialised as being uncertain, and at each time step each agent randomly selects a behaviour type according to probabilities given in table 3. Furthermore, at each time step each agent fuses their current truth-value with probability σ and no fusion takes place for that agent with probability $1-\sigma$. In the case that fusion occurs another agent is selected at random to act as the transmitting agent. Also, at each time step each agent decides whether or not to look for evidence according to their current behaviour type (see figure 1). If an agent searches then they receive evidence with probability ρ. For all combinations of parameter settings the results are averaged over 100 independent runs, each running for 1500 time-steps. Error bars then correspond to 90% percentiles for the data obtained from the independent runs.

Figure 9 shows the proportion of agents with truth-values **t** (black line), **u** (green line) and **f** (blue line) plotted against time, for three mixtures of behaviour types. The results are for $\sigma = 0.9$ and $\rho = 0.1$ which is a challenging scenario when the environment is dynamic since direct evidence is relatively scarce compared to the frequency of agent interactions, making it more difficult for the population to detect an underlying change in the state of the world. Figure 9a is for a homogeneous population consisting only of **AC** agents. In this case, the agents quickly reach consensus on what is initially the true state-of-the-world, i.e. **f**, but given the subsequent dominance of this truth value in the population, agents are then unable to learn the new state of the world after the change at $t = 500$. There is similar behaviour shown in figure 9b which is for a $50/50$ mixture of **AC** and **CC** behaviours. In this case it takes the population slightly longer to reach consensus on the initial correct state **f**, but again agents are unable to revise their truth-values at $t = 500$. We hypothesise that in both figures the population reaches consensus close to the proportions $(1, 0, 0)$ in the time period up to $t = 500$. This is an equilibrium point for both mixtures of behaviour types making subsequent updating impossible. In contrast, figure 9c show results for $\lambda = 0.01$ and $w = 0.5$ where there is a $50/50$ split between adventurous and cautious fusion behaviours, and where there is a small but non-zero proportion of inquisitive evidential updating behaviour in a population otherwise dominated by the confident behaviour type. In this case, we see that the population initially reaches consensus on **f**, but at $t = 500$ we see a rise in the proportion of **u** agents followed by a gradual increase in the proportion of **t** agents until consensus is reached on the new state-of-the-world by $t = 1500$. In the first phase of learning for $t < 500$ figures 9c and 9b show almost identical results, but the small proportion of inquisitive agents that continue to collect evidence throughout are sufficient to allow the population to adapt when the state-of-the-world changes.

Conclusions & Future Work

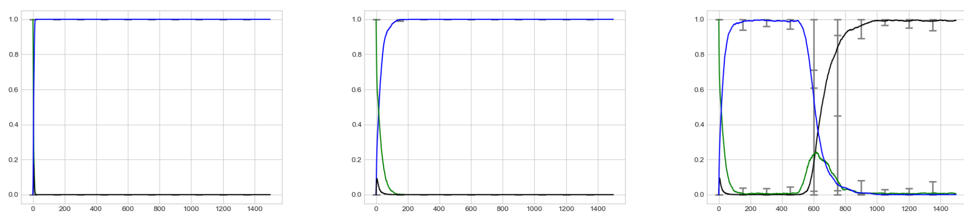
We have presented social learning as a combination of two processes; fusion in which agents change their beliefs under the influence of their peers and evidential updating in which agents learn directly from evidence. In this context we have considered four overall behavioural types generated by independently combining conservative and open-minded approaches to both processes. Different behaviour types have been shown to have different convergence and consensus properties. However, certain heterogeneous mixtures perform best in a range of different learning scenarios. In particular, a $50/50$ mix of adventurous and cautious fusion combined with a mix of 1% inquisitive and 99% confident evidential updating, is highly robust especially in dynamic environments.

In this paper we have assumed a fully connected, well-mixed population of interacting individuals. This is a strong assumption and there is growing evidence that it is not always optimal for social learning (Crosscombe & Lawry, 2022). Future work will investigate social networks with more constrained topologies in which connection between heterogeneous behaviour types may vary. Following



(a) $\sigma = 0.1$ and $\rho = 0.1$ (b) $\sigma = 0.9$ and $\rho = 0.1$ (c) $\sigma = 0.1$ and $\rho = 0.9$ (d) $\sigma = 0.9$ and $\rho = 0.9$

Figure 8: Proportion of correct beliefs with different combinations of parameter values w and λ and for different fusion and evidence rates. Results are for initial proportions $(1, 0, 0)$, i.e. all agents are wrong, $\epsilon = 0.3$ and at $t = 1500$



(a) $\lambda = 0, w = 1$ (b) $\lambda = 0, w = 0.5$ (c) $\lambda = 0.01, w = 0.5$

Figure 9: Average proportion of $\mathbf{t}$ (black line), $\mathbf{u}$ (green line) and $\mathbf{f}$ (blue time) belief states, plotted against time for different mixtures of behaviour types. Results are for 200 agents, $\sigma = 0.9$, $\rho = 0.1$ and $\epsilon = 0.3$ and are averages over 100 independent runs. The true state of the world switches from $\mathbf{f}$ to $\mathbf{t}$ at $t = 500$.

(Erbach-Schoenberg et al, 2011) it may also be interesting to consider heterogeneous updating rates e.g. where uncertain individuals update more often or more readily. Taking these ideas forward we will then explore their direct application to modelling social learning in biological and artificial systems.

Acknowledgements

This work was funded and delivered in partnership between Thales Group, University of Bristol and with the support of the UK Engineering and Physical Sciences Research Council, ref. EP/R004757/1 entitled “Thales-Bristol Partnership in Hybrid Autonomous Systems Engineering (T-B PHASE).”

References

A. Baronchelli: *The emergence of consensus: a primer*. R. Soc. open sci.5:172189. 2018

M. Brambilla, E. Ferrante, M. Birattari, M. Dorigo: *Swarm robotics: a review from the swarm engineering perspective*. Swarm Intelligence. 7. 1–41. 2013

E. N. Buckingham: *Division of labor among ants*. Proc. Am. Acad. Arts Sci. 46. 425-508. 1911

M. Crosscombe, J. Lawry: *Exploiting vagueness for multi-agent consensus*. Multi-Agent and Complex Systems. 67-78. Springer. 2017

M. Crosscombe, J. Lawry, S. Hauert, M. Homer: *Robust distributed decision-making in robot swarms: Exploiting a third truth state*. IEEE/RSJ International Conference on Intelligent Robots and Systems. 4326-4332. IEEE. 2017

M. Crosscombe , J. Lawry: *The Impact of Network Connectivity on Collective Learning*. Distributed Autonomous Robotic Systems. DARS 2021. Springer Proceedings in Advanced Robotics. 22. Springer. 2022

I. D. Couzin, C. C. Ioannou, G. Demirel, T. Gross, C. J. Torney, A. Hartnett, L. Conradt, S. A. Levin, N. E. Leonard: *Uninformed Individuals Promote Democratic Consensus in Animal Groups*. Science. 334. 1578-1580. 2011

M. Črepinšek, S. H. Liu, M. Mernik: *Exploration and exploitation in evolutionary algorithms: A survey*. ACM Comput. Surv. 45. 3(35). 2013

I. Douven, C. Kelp: *Truth Approximation, Social Epistemology, and Opinion Dynamics*. Erkenntnis. 75. 271. 2011

E. Z. Erbach-Schoenberg, C. McCabe, S. Bullock: *On the interaction of adaptive timescales on networks*. Proceedings of the Eleventh European Conference on Artificial Life. 900-907. MIT Press. 2011.

G. Fu, W. Zhang: *Opinion Dynamics of Modified Hegselmann-Krause Model with Group-based Bounded Confidence*. Proceedings of the 19th World Congress The International Federation of Automatic Control. 9870-9874

C. M. Heyes: *Social learning in animals: categories and mechanisms*. Biol. Rev. 69. 207-231. 1994

E. R. Hunt, B. Mi, R. Geremew, C. Fernandez, B. M. Wong, J. N. Pruitt, N. Pinter-Wollman. *Resting networks and personality predict attack speed in social spiders*. Behavioral Ecology and Sociobiology. 73. 97. 2019

T. A. O'Shea-Wheller, E. R. Hunt, T. Sasaki: *Functional Heterogeneity in Superorganisms: Emerging Trends and Concepts*. Annals of the Entomological Society of America. XX(X). 1–13. 2020

C. A. C. Parker, H. Zhang: *Cooperative decision-making in decentralized multiple-robot systems: The best-of-n problem*. IEEE/ASME Transactions on Mechatronics. 14(2). 240–251. 2009

J. M. Peters, O. Peleg, L. Mahadevan: *Collective ventilation in honeybee nests*. J. R. Soc. Interface 16:20180561. 2019

J. Prasetyo, G. De Masi, · E. Ferrante: *Collective decision making in dynamic environments*. Swarm Intelligence. 13. 217–243. 2019

L. Schäfer, F. Christianos, J. Hanna, S. V. Albrecht. *Decoupling exploration and exploitation in reinforcement learning*. Unsupervised Reinforcement Learning (URL) Workshop in the 38th International Conference on Machine Learning.

A. Slivkins. *Introduction to Multi-Armed Bandits*. Foundations and Trends in Machine Learning. 12(1-2). 2019

A. Smith: *An inquiry into the nature and causes of the wealth of nations*. W. Strahan and T. Cadell, London. 1776

G. Theraulaz, E. Bonabeau, J-L. Deneubourg. *Response threshold reinforcement and division of labour in insect societies*. Proc. R. Soc. B. 265. 327-332. 1998

G. Valentini, E. Ferrante, M. Dorigo: *The Best-of-n Problem in Robot Swarms: Formalization, State of the Art, and Novel Perspectives* Front. Robot. AI. 4(9). 2017

A. Yaman, N. Bredeche, O. Caylak, J. Z. Leibo, S. Wan Lee: *Metacontrol of Social Learning Strategies*. PLOS Computational Biology, 18(2). 2022