



Fox, R., & Bullock, S. (2022). Nectar of the Bots: Evolving Bidirectional Referential Communication. *Adaptive Behavior*.
<https://doi.org/10.1177/10597123221110379>

Publisher's PDF, also known as Version of record

License (if available):
CC BY

Link to published version (if available):
[10.1177/10597123221110379](https://doi.org/10.1177/10597123221110379)

[Link to publication record in Explore Bristol Research](#)
PDF-document

This is the final published version of the article (version of record). It first appeared online via Sage at <https://doi.org/10.1177/10597123221110379>. Please refer to any applicable terms of use of the publisher.

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:
<http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/>

Nectar of the Bots: Evolving Bidirectional Referential Communication

Ruairi Fox¹ and Seth Bullock¹ 

Adaptive Behavior
2022, Vol. 0(0) 1–22
© The Author(s) 2022



Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/10597123221110379
journals.sagepub.com/home/adb



Abstract

Referential communication is central to social and collective behaviour, for example honey bees communicating nectar locations to each other or co-workers gossiping about a colleague. Since such behaviour typically is considered to be ‘representation hungry’, it is often assumed to require the possession of complex cognitive machinery capable of manipulating symbolic representations of the world. However, a series of simulation studies have shown that it can be achieved by very simple embodied artificial agents controlled by evolved recurrent artificial neural networks that are challenging to interpret in symbol-processing terms. In this paper, we extend this paradigm to explore scenarios in which a pair of agents, each of which is privy to a different piece of private information, must jointly solve a task that requires *both* pieces of information to be communicated, compared and acted upon, i.e., each agent must simultaneously play the role of both signaller and receiver during an unstructured referential communication interaction that is *bidirectional*. We demonstrate evolved agents that are able to solve this task, and analyse the extent to which their situated, embedded and embodied communicative behaviour can be considered to be a step towards understanding the minimal cognitive basis for human language.

Keywords

Communication, neural networks, evolution, agents

Introduction

It is commonplace, or arguably even ubiquitous, for people and other animals to use signalling behaviours to convey referential information. That is, signallers often use signals to inform observers about hidden, private or remote states of affairs that the observers would otherwise be ignorant of.

One classic example is the behaviour of group-living honey bees, which communicate the location of resources such as nectar-rich flowers and potential nest sites by performing a ‘waggle dance’ for their nest-mates (von Frisch, 1965). During this dance, a bee will waggle back and forth energetically whilst repeatedly tracing a characteristic figure-of-eight path on a ‘dance floor’ within the hive. Watching bees are able to infer the direction, distance and quality of the signalled resource by attending to the orientation, duration and vigour of the figure-of-eight dance.

Whilst the complexity of this bee communication system is extremely impressive given their relatively small brains, human language stands as the most sophisticated vehicle for referential communication so far discovered. Human languages differ from the honey bee *Tanzsprache* in many ways. Perhaps most distinctively, they exhibit complex, compositional syntax and can be used to achieve a very

wide range of functions, from giving orders and making promises to reciting poems and telling jokes. However, like the bee waggle dance, the primary function of human language is referential communication, allowing one individual to inform another of something that they did not already know.

Whether in humans or less sophisticated creatures, referential communication is often assumed to involve complex cognitive machinery capable of building and processing internal symbolic representations of the world, and translating these structure-sensitive, quasi-linguistic representations into external, meaningful symbols that can be decoded by a suitably sophisticated observer (Fodor & Pylyshyn, 1988). However, a series of studies have demonstrated that this type of communication can in fact be

¹Computer Science, University of Bristol, Bristol, UK

Handling Editor: Eduardo Izquierdo

Corresponding author:

Seth Bullock, Computer Science, University of Bristol, 70 Woodland Road, Bristol BS8 1UB, UK.

Email: seth.bullock@bristol.ac.uk

achieved in the absence of any recognisable ‘language of thought’ (Fodor, 1975) by very simple simulated agents that possess no pre-defined language, no pre-established lexicon of available symbols, no pre-arranged, dedicated signalling channel, and no pre-set notion of turn taking or shared attention (Campos & Froese, 2017, 2019; Quinn, 2001; Williams et al., 2008). These agents are situated in simple environments and are typically controlled by small Continuous Time Recurrent Neural Networks (CTRNNs; Beer, 1995) that are not easy to interpret in symbol-processing terms (Manicka, 2012).

To date, these studies have tended to consider scenarios in which one signaller agent is tasked with informing one receiver agent about one aspect of the environment. Either the first agent employs a dedicated signaller strategy whilst the second employs a separate dedicated receiver strategy (Williams et al., 2008), or both agents use the same strategy but are each primed to play either the signaller role *or* the receiver role during a particular signalling episode (Campos & Froese, 2017, 2019). By contrast, real-world referential communication is often *bidirectional*, with each agent attempting to both inform and be informed by the other during the same signalling episode. A pair of competing stags might each communicate their resource holding potential (fighting ability) to the other in advance of fighting (Parker, 1974), or two friends might debate which pub to visit.

In this paper, we extend the paradigm employed in the studies mentioned above in order to explore scenarios in which a pair of agents, each of which is privy to a different piece of private information, must jointly solve a task that requires both pieces of information to be shared, compared and acted on, i.e., each agent must play the role of both signaller and receiver in an unstructured bidirectional referential communication interaction. We demonstrate conditions under which evolved agents are able to solve this task, and analyse the extent to which their situated, embedded and embodied communicative behaviour can be considered to be truly ‘symbolic’, before considering the extent to which this line of research has the potential to shed light on the minimal cognitive basis for communication that is as complex as human language.

Previous Studies

Defining Referential Communication

Communication has proven somewhat difficult to define within the animal behaviour community and also the adaptive behaviour community. On the one hand, defining communication as any episode in which one agent influences another one is too loose, allowing pushing and shoving to count as communication. On the other hand, requiring that communication must involve signallers encoding conceptual content in syntactically structured signals

composed from meaningful symbols that are then decoded by an observer in order to create or update their internal representation of the world presupposes too much about the symbolic nature of communication and allows no room within the category for proto-languages and other precursors to full-blown human language.

Moreover, whilst requiring that a signalling interaction must increase the fitness of the signal *receiver* (Johnstone, 1997) excludes interactions that are malicious, deceptive or manipulative, it also rules out genuine efforts at communication that happen to be mistaken, redundant or ineffective in some way. Similarly, requiring that the signal *producer* gains a direct fitness benefit from signalling may exclude self-sacrificing communications that benefit the receiver(s) to the detriment of the signaller, for example deliberate admissions of guilt, as when a criminal admits to their crime.

Even requiring, as Maynard Smith and Harper (2003, p.15) do that a signal is ‘an act or structure that alters the behavior of another organism, which evolved because of that effect, and which is effective because the receiver’s response has also evolved’ leaves unspecified what the receiver’s signal consuming behaviour evolved *for*, opening the door to the inclusion of passive mimesis and masquerade behaviours as types of communication (as when an insect species has evolved to resemble a twig in order to avoid being predated).

For the purposes of this paper, we will propose and employ the following definition:

Referential communication occurs when the signal-producing behaviour of one agent (the *signaller*) has the proper function to adapt a second agent (the *receiver*), via its sense organs, to some state of affairs, and when this second agent’s signal-consuming behaviour has the proper function to be so adapted.

Here, the ‘proper function’ of an agent behaviour should be understood in the historical (teleosemantic) sense established by Millikan (1989b): roughly, the proper function of an evolved device, *D*, is to perform the function that, when performed by *D*’s ancestors, led to proliferation of the genes responsible for *D*’s existence. Thus, the function of a heart is to pump blood, because it was by pumping blood, rather than being red and wet, or making a bumpety-bump noise, that ancestral hearts contributed to the proliferation of genes for hearts.

Proper functions are *normative*. For example an injured heart that cannot pump blood is *malfunctioning* – it has the function to pump blood (by virtue of its evolutionary history), but is currently unable to carry out this function. Note also that proper functions are not straightforwardly *statistical* or *causal* dispositions: a sperm has the function of fertilising an egg (by virtue of the evolutionary history of sperm) regardless of the fact that the vast majority of sperm

that have existed have not managed to carry out this function.

Notice that whilst key signalling scenarios fall *within* the definition presented above, some other superficially similar types of interaction do not. According to the above definition, for example, when one agent makes an alarm call in order to alert another agent to the fact that a predator is approaching, and, in response, the second agent instinctively hides or flees, this is a canonical example of successful referential communication. However, since the definition does not specify that communication behaviours must succeed, it also allows *failed* signalling attempts to count as referential communication. For example the following would all count as faulty, defective or malfunctioning signalling behaviours: if no receivers are present or able to hear the signaller's alarm call, or the signaller makes the alarm call in error when their belief that a predator is approaching is mistaken, or the signaller tries to make the alarm call but makes a different call by mistake, or a gust of wind blowing through a hollow log sounds to a receiver like an alarm call and causes them to run and hide unnecessarily.

By contrast, the definition rules out scenarios that resemble signalling but, upon closer examination, are more accurately described in terms of 'mind-reading' or 'manipulation' (Krebs & Dawkins, 1984) in which one agent exploits the 'tacit suppositions' or 'behavioural biases' of another (e.g. Bullock, 1998). For instance if one agent gives away some information to another agent *accidentally*, as when a poker player reveals that they are bluffing by subconsciously making an involuntary 'tell' such as rubbing their nose, this will *not* count as referential communication because the unfortunate player's informative behaviour does not have the *proper function* of adapting the observer's behaviour to the fact that they have a weak poker hand (Millikan, 1984). Conversely, if one agent *deliberately misleads* another by, for example giving the alarm call that triggers in receivers an instinctive flee response associated with predator attack, when it knows that there is in fact no predator, this will also *not* count as referential communication (although this behaviour is parasitic on an existing referential communication system) since, although the signaller's behaviour functions to adapt the behaviour of the listening agent to a state of affairs (that a predator is approaching), the deceived receiver's behaviour does not have the proper function of being so adapted under circumstances in which there is in fact no predator approaching (Artiga, 2014; Bullock, 1997; Noble et al., 2001).

Note that the above definition does not require that the state of affairs being communicated about needs to be remote in space or time from either the signaller or the receiver. A courtroom witness indicating the identity of one criminal by pointing at them with her finger and then naming a second criminal who was still on the run would count as engaging in two instances of referential

communication. However, the *displaced reference* that is employed in the second example is a powerful feature of human communication and is understood to otherwise be rare throughout natural signalling systems. For example, whilst the honey bee dance language can handle *spatial* displacement to some extent, it is not clear that honey bees can communicate about *temporally* displaced states of affairs in the past or future.

Finally, note that whilst this framing of signalling is consistent with the notion that referential communication involves the exchange of quasi-linguistic symbolic representations within which signallers encode information about the state of the world that can subsequently be decoded by receivers, this is not an explicit requirement of the definition (or of the teleosemantic approach to language and mental content in general) which remains neutral about the exact nature of the signal producing and consuming devices and the precise mechanisms by which agents use communication to adapt each other to states of affairs. Consequently, a dynamical systems theoretic interpretation of cognitive behaviour (Beer, 2000) is equally consistent with this model of communication.

Evolving Referential Communication

Perhaps the most seminal paper in this area was published by Quinn (2001), who demonstrated the evolution of successful communication behaviour in a pair of simulated model agents that were not already provided with a dedicated communication channel or a set of symbols with which to communicate. The context for Quinn's work was a number of prior papers that claimed to have demonstrated the evolution of language in simple simulated agents (e.g. MacLennan & Burghardt, 1993; Werner & Dyer, 1992).

One study that exemplifies this style of work is due to Werner and Dyer (1992) who evolved a pair of simulated agents to solve a 'mating' problem within a 200-by-200 grid world. In this scenario, one agent was mobile and capable of receiving 'acoustic' signals but was otherwise 'blind', whilst a second agent was 'sighted' and capable of producing 'auditory' signals but remained stationary. The joint task of the pair of agents was to successfully navigate the mobile agent to the cell occupied by the stationary agent. At each time step, the stationary agent used its evolved rule-set and its visual appraisal of the current location of the mobile agent to determine which of its repertoire of auditory signals it should produce. Each signal was represented by a three-bit string, allowing for eight distinct sounds to be employed. Simultaneously, at each time step the mobile agent used its evolved rule-set plus the auditory signal that it received from the stationary agent to determine in which direction it should move. After many generations of evolution, the agents co-ordinated on rule-sets that allowed the mobile agent to consistently navigate to the location of the

stationary agent – despite the relationship between the available sounds and the available movement behaviours being entirely arbitrary, initially random and evolutionarily unconstrained. The results of the paper showed that a successful communication scheme could arise spontaneously within a simple agent system, and that agents capable of communication were able to solve the mating problem in half as many moves as agents that were evolved without the ability to communicate.

However, whilst this style of work was extremely influential, the extent to which it could claim to shed light on the evolution of language was limited by the existence within the model of dedicated communication channels and discrete sets of pre-defined symbols (e.g. eight different sounds) and pre-defined actions (e.g. taking a step in one of eight different directions), which ensured that the challenge facing agents was not to evolve a language from scratch (whatever that might mean) but rather to simply co-ordinate on an appropriate lexicon that maps eight different symbols to eight different behaviours (Steels, 1997). A different approach would be needed if models were to help explain either the dynamics of truly grammatical languages (Kirby, 2002) or the cognitive basis for communication itself (Quinn, 2001).

In response, Quinn (2001) introduced an evolutionary simulation model (Bullock, 1997) in which the communicating agents were idealised versions of small, wheeled Khepera robots, possessing eight infra-red sensors (each with a 5 cm range) and two independently motor-driven wheels capable of achieving a maximum speed of 8 cm per simulated second (see Figure 1). Pairs of these agents, each controlled by the same evolved artificial neural network, were initially placed close to each other but oriented at random, within a continuous infinite 2D space, and were jointly tasked with travelling at least 25 cm (i.e., 10 agent radii) from their starting location in any direction within 10 simulated seconds whilst remaining within sensor range of each other and without colliding. This task was designed such that, in order to succeed, both agents needed to agree on the same direction of travel without recourse to any compass, landmarks, etc.

A successful strategy was evolved in which the pair of agents dynamically co-allocate the roles of ‘leader’ and ‘follower’ (Figure 2). This is achieved by each agent first rotating until its infra-red sensors indicate that it is facing its partner. The agent that manages to do so first takes the role of follower and moves back and forth until the other agent has turned to face it. At this point, the follower approaches the leader and the leader retreats backwards (whilst both maintain a fixed separation distance), resulting in both moving off together in an arbitrary but consistent direction.

By contrast with previous papers, in this case the agents could not be described as having solved the communication task by agreeing on how to use a pre-defined vocabulary of

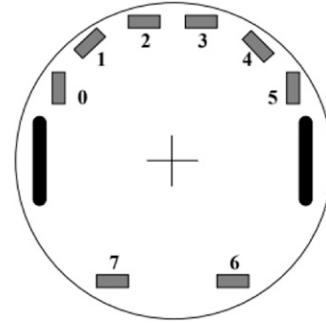


Figure 1. The positions of sensors (numbered grey rectangles) and actuators (black lozenges) on a Khepera robot. Figure taken from Quinn (2001).

symbols. It was not even clear that the communication strategy involved any symbol use at all, or relied on the use of any internal representations of aspects of the world. In fact, the evolutionary history of the strategy reveals that initially non-communicative behaviours that allowed one agent to find and move towards the other agent without colliding with it had become ritualised into a dance-like behaviour that allowed the agents to co-ordinate their movement in order to solve the task. Rather than lending itself to a quasi-linguistic symbol-processing interpretation, the evolved behaviour was more readily explicable in terms of dynamical systems ideas of coupling and synchronisation.

However, this aspect of the work can also be considered a shortcoming. Was the evolved behaviour in fact non-communicative, being an example of mere *coordination*? Even if the evolved behaviour could be classed as communication, what exactly was being communicated? The evolved behaviour was not a clear example of *referential* communication, and was certainly not *displaced* referential communication, as the task did not require that the agents engage with anything beyond their immediate sensor readings.

In an attempt to demonstrate the evolution of artificial communication conclusively, Williams et al. (2008) evolved successful agent communication in the context of a more complex task – one that could only be solved by displaced referential communication. Here, one agent (known as the sender) must communicate a target location to a second agent (known as the receiver), and the receiver must then navigate to this location. The agents operated within a 1D periodic environment (a ring) and were each controlled by a five-neuron CTRNN. Each agent possessed two proximity sensors, one extending clockwise and the other anticlockwise, each reporting the angular distance to the other agent to a range of $\pi/8$. Each agent also possessed a bearing sensor that either indicated their own current angular location (for the receiver) or the angular separation between their current

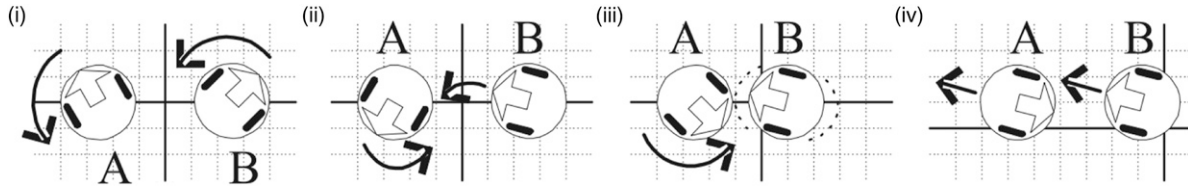


Figure 2. Communication behaviour exhibited by the pair of agents evolved by Quinn (2001): (i) each initially randomly oriented agent rotates until it senses its partner; (ii) B senses A first; (iii) B oscillates back and forth until A senses B; (iv) A and B move off together in the direction indicated by the orientation of the agent that was first to detect its partner (B in this case). Figure taken from Quinn (2001).

position and the target location (for the sender). A diagram of the environment and the agents is provided in Figure 3.

Each individual communication strategy was determined by the parameters of two CTRNNs, a signaller network and a receiver network, but both were encoded on the same genome ensuring that one evolutionary individual represented one entire solution to the problem, composed of an explicit sender strategy plus a separate explicit receiver strategy. Some form of communication is necessary in order to solve the task as the receiver must navigate to a target location about which they initially are ignorant. In the first experiment, no constraints were placed on agent interaction which led to the evolution of two main strategies: ‘shepherding’ and ‘sit and wait’. In the former case, the sender would ‘push’ and ‘pull’ the receiver in order to guide them to the goal, whereas in the latter the sender would stop within proximity sensor range of the target in order to indicate its location. However, the fact that the agents solved the task whilst remaining within proximity sensor range of each other ensured that whilst the evolved behaviour was clearly referential communication it could not be categorised as *displaced* referential communication.

To remedy this, in the next experiment the sender’s movement was restricted such that it was constrained to remain within one quarter of the environment, with the targets located in one of four locations outside of this region (see Figure 3). This restriction renders the previously evolved strategies ineffective and instead requires the use of displaced referential communication. Evolved agents were able to succeed at the task, but did not generalise well to target positions beyond the four that they were evolved to deal with. The evolved solution effectively relied on a set of four distinct movement signals, each associated with one of the four target locations, but had little ability to communicate about points lying between these locations. This evolved communication system is described by the authors as ‘symbolic’ and is compared to the use of a simple set of ‘words’. In a final experiment, the number of target locations was increased to 10 (Figure 3), encouraging the evolution of a communication strategy that could generalise over the entire range of target locations, demonstrating that

it was possible to evolve displaced generalised referential communication in simple agents.

Campos and Froese (2017) extended this study by simplifying the agents involved in order to allow for more analysis to be performed, and to demonstrate that signalling and receiving behaviour could be handled by the same 3-node CTRNN. They employed an infinite non-periodic 1D environment (a line), with one target appearing uniformly at random in the range $[0.5, 1]$ and two agents initially placed independently at locations drawn uniformly at random from the range $[0, 0.3]$ (see Figure 4, top). The sender agent could sense the location of the target but was not permitted to leave the ‘signalling zone’ $[0, 0.3]$, whilst the receiver could not sense the location of the target but was permitted to move freely throughout the environment. Each agent had three sensors: a binary contact sensor that was on only if agents were separated by a distance of less than 0.4, a self-position sensor that indicated the agent’s own location in the world and a target sensor. For the sender, the target sensor indicated its current distance to the goal. For the receiver, this sensor gave a constant reading of -1 . This allowed for the receiver and sender to behave differently during a trial, despite being controlled by networks with the same parameters. The evolved agents were able to successfully communicate the location of the target over the range $[0.5, 1]$ using a displaced referential communication strategy in which the location of the target was related to the amount of time that the agents spent within contact range.

Subsequently, the authors were able to extend this model to a 2D environment within which a pair of agents controlled by the same evolved 6-neuron CTRNN were able to communicate the location of a target using a displaced referential communication scheme that made use of both the bearing and distance to a target location in a compositional manner with parallels to the honey bee’s waggle dance language (Campos & Froese, 2019). This paper represents the state of the art in this line of research, but leaves open the challenge of exploring whether two-way communication can be achieved by similar cognitive architecture and also the question of to what extent these studies and studies like them can shed light on the evolution of more sophisticated linguistic behaviour.

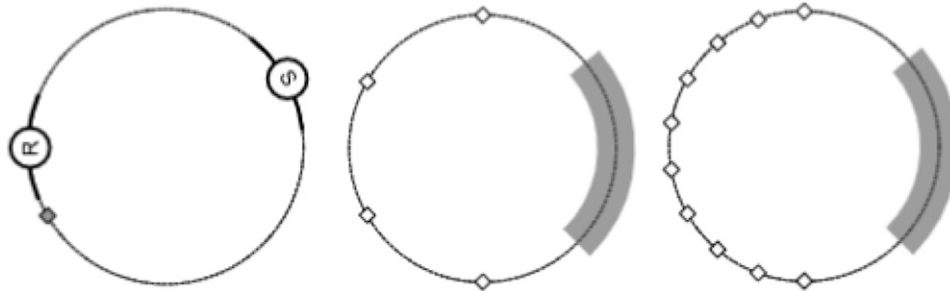


Figure 3. The environment and agents employed by Williams et al. (2008). *Left:* The sender is marked with an S and the receiver with an R. A stationary goal (grey diamond) is randomly positioned at the start of each trial. *Centre:* The region that the sender was constrained to in the second set of experiments (grey) and the four possible target locations (diamonds). *Right:* The region that the sender was constrained to in the third set of experiments (grey) and the 10 possible target locations (diamonds). Figures taken from Williams et al. (2008).

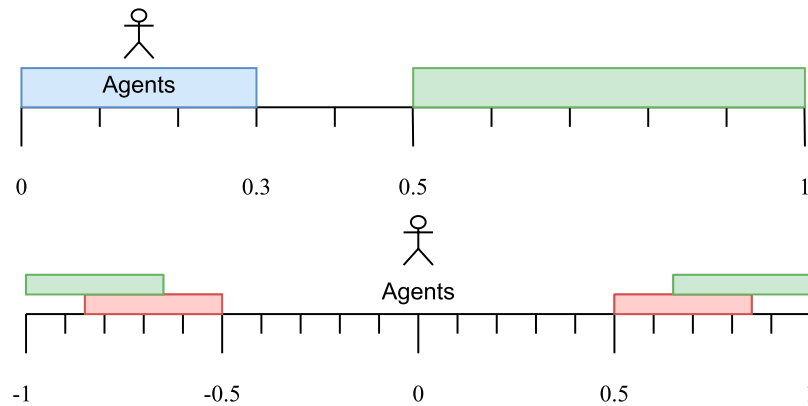


Figure 4. *Top:* The one-way communication task environment. Both agents are initially randomly located within the $[0, 0.3]$ region. The receiver may move freely, but the signaller is restricted to remain within this region. For each trial, the target is randomly located within the region $[0.5, 1]$. *Bottom:* The two-way communication task environment. Both agents are initially located at the origin and both may move freely. The true target is randomly located within one of the upper (green) regions $\pm[0.65, 1]$. The false target is randomly located within one of the lower (red) regions $\pm[0.5, 0.85]$. The values of the target coordinates are guaranteed to differ in sign, and the absolute values of the target coordinates are also guaranteed to differ by at least 0.15 distance units.

Evolving Bidirectional Referential Communication

In this paper we will consider two different referential communication tasks. The first, which we term the *one-way communication task*, is taken from Campos and Froese (2017). The second, bidirectional task, which we will term the *two-way communication task*, is defined here for the first time.

One-Way Communication

During the one-way communication task, two agents move and interact within a simple 1D environment for 300 simulated units of time, one playing the role of *signaller* and the other playing the role of *receiver*. The signaller can sense

the location of a stationary target within the environment and aims to communicate this information to the receiver. The receiver is not able to sense the location of the target but aims to navigate to it based on information that it obtains from the signaller. The receiver and signaller can each detect their own location and can also detect whether or not they are within a threshold distance of each other, but have no other means of influencing one another.

At the start of each trial, both agents have their neuron activation levels set to zero and their positions initialised to locations selected uniformly at random from the range $[0, 0.3]$, and the target location, G , is placed at a location drawn uniformly at random from the range $[0.5, 1]$. During a trial, the signaller's movement is restricted to lie within a 'signalling zone' $[0, 0.3]$, whereas the movement of the receiver is unrestricted (see Figure 4). Each agent

Table 1. Sensor Definitions for Sensors Employed in the One-way Communication Task and the Two-way Communication Task.

One-way communication task:	
1. Self-Position sensor	A continuous-valued sensor that returns the position of the agent in the environment
2a. Distance to target sensor ...or 2b. Constant value sensor	A continuous-valued sensor that returns the distance between the agent and a target location A sensor that always returns the value -1
3. Contact sensor	A binary sensor that returns $+1$ if the distance to the other agent is less than or equal to 0.4 , otherwise 0
Two-way communication task:	
1. Self-position sensor	A continuous-valued sensor that returns the position of the agent in the environment
2. Distance to target sensor	A continuous-valued sensor that returns the distance between the agent and a target location
3. Proximity sensor	A continuous-valued sensor that returns the distance separating the two agents

receives input from its three sensors, each providing an external weighted input to one unique node within its own 3-node CTRNN controller (see Table 1). Firstly, each agent possesses a binary *contact sensor* which delivers a value of $+1$ if the two agents are separated by a distance of less than 0.4 and a value of 0 otherwise. Secondly, each agent also receives input from its *self-position sensor*, which delivers a value corresponding to its own real-valued co-ordinate within the environment. The agents differ in that their third sensor either provides their distance to the target location (if the agent is the signaller) or provides a constant value of -1 (if the agent is the receiver).

Each agent's behaviour is controlled by a small Continuous Time Recurrent Neural Network (CTRNN) comprising three model neurons that are updated by Euler integration of the standard equation for a CTRNN neuron's dynamics:

$$\tau_i \frac{dy_i}{dt} = -y_i + \sigma\left(\beta_i + \sum w_{ij}y_j + w_{iI}I_i\right) \quad (1)$$

Here, τ_i is the time constant of neuron i , y_i is the current activation of neuron i , β_i is the bias of neuron i , w_{ij} is the weight on the connection from neuron j to neuron i , I_i is the magnitude of any external sensory input to node i , w_{iI} is the weight on this sensory input channel, and $\sigma(\cdot)$ is the neuron activation function, which is the standard sigmoid function: $\sigma(x) = (1 + e^{-x})^{-1}$

Each agent's behaviour is controlled by a separate copy of the *same* genetically encoded CTRNN. That is the CTRNNs controlling the signaller and the receiver have the same structure (i.e. weights, biases and time constants), but each receives its own sensory inputs and maintains its own internal state. All neuron activation values are set to zero at the start of each trial and each simulation step represents 1 unit of simulated time.

Each agent's movement is driven by the activation level of one CTRNN node. The maximum speed of each agent is fixed at 0.01 units of distance per unit of simulated time. (Acceleration, momentum, friction, etc., are not considered as part of the model.) For each simulation step, each agent's

movement is updated using equation (2) below, where Δ_x is the change in the agent's position, $\Delta_t = 1$ is the number of units of time that pass during one simulation step (which is small with respect to the range of neuron time constant values), $V = 0.01$ is the maximum speed at which an agent can travel, and M is the activation of the neuron that drives the motor, which lies in the range $[0, 1]$. Note that whereas each sensory input to the network is weighted, there is no weight on the output neuron's connection to the motor.

$$\Delta_x = 2(M - 0.5)V\Delta_t \quad (2)$$

Two-Way Communication

The two-way communication task is designed to be as comparable with the original one-way communication task as possible, but differs in several key respects.

During the two-way communication task, two agents move and interact within a simple 1D environment, each playing the role of both *signaller* and *receiver*. One of the agents is able to directly sense the location of one randomly located stationary target within the environment, whilst the other agent is able to directly sense the location of a second randomly located stationary target. The target that is furthest from the origin is the *true* target for both agents. The other target is an irrelevant *false* target. The target locations are generated such that (i) they lie on opposite sides of the origin, (ii) they each lie between 0.5 and 1.0 distance units from the origin and (iii) the difference between the absolute values of their locations is at least 0.15 , meaning that one is significantly further from the origin than the other.

Despite each agent not knowing whether they can sense the true target or the false target, their joint aim is for both of them to navigate to the true target, the one that is furthest from the origin. To do this they must exchange information in order to (i) determine which target to approach, and (ii) (for one of the agents) determine where this target is.

At the start of each trial, both agents are located at the origin with neuron activation levels set to zero. The location of the *true* target, G , is chosen uniformly at random from either the range $[0.65, 1]$ or the range $[-0.65, -1.0]$, with equal frequency. If the true target, G , is positive, a second *false* target location, g , is chosen uniformly at random from a range of negative locations $[-0.5, -(G - 0.15)]$, otherwise g is chosen uniformly at random from a range of positive locations $[+0.5, -(G + 0.15)]$, thus, ensuring that $|G| - |g| \geq 0.15$, i.e., the absolute location of the true target is at least 0.15 greater than the absolute location of the false target.

During a trial, the movement of both agents is unrestricted (see Figure 4). Each agent again receives input from three sensors, each providing an external weighted input to one unique node within its own 3-node CTRNN. Firstly, each agent possesses a sensor which indicates the signed distance separating the agents from each other (i.e. the contact sensors from the original one-way communication task have been replaced with *proximity sensors*). Secondly, each agent receives input from its *self-position sensor*. Finally, each agent's third sensor now provides the distance to one unique target location; one agent senses the distance to the location of the true target whilst the other senses the distance to the location of the false target. The agents are not aware of whether they are sensing the true target or the false target.

The scenario is otherwise identical to the one-way communication task.

Evolving CTRNNs

For each of the two tasks described above, a genetic algorithm (GA) was employed to discover successful solutions. The scheme employed is based on that described by Campos and Froese (2017). Each evolutionary generation comprised a population of $N = 50$ individual genomes, each a vector of 18 real values in the range $[-1, +1]$ used to parameterise a single CTRNN. The 18 values encoded three sensory input weights (from the contact/proximity sensor, the self-position sensor and the target sensor, respectively), and a bias, time constant and three inter-neuron weights for each of the three CTRNN neurons. Each gene value was linearly re-scaled to a range appropriate to the phenotypic component that it encoded: weights and biases were mapped to the range $[-16, +16]$ and time constants to the range $[50, 100]$. Note that one individual genome is used to encode the parameters of both agents' CTRNNs, which share the same weights, biases, time constants and initial activation values.

For the one-way communication task, the score for each individual trial was calculated on the basis of the distance between the target location and the final position of the receiver agent as follows:

$$S_i^T = \max(0, 1 - |R_i^T - G^T|) \quad (3)$$

Where S_i^T is the score that was achieved on trial T by individual i from the current population of networks, R_i^T is the location of the simulated receiver agent controlled by network i at the final time step of trial T , and G^T is the location of the target during trial T .

For each individual network, fitness was calculated on the basis of performance across 20 trials, each randomly specifying the location of the target and the initial locations of the two agents. For each generation, the same 20 trials were used to evaluate each network in the population. The fitness of an individual network was calculated as a weighted sum of the scores that it achieved across these 20 trials, where each score's weight was equal to the reciprocal of its rank (ascending) within the list of 20 trial scores for that individual, thus, ensuring that the individual's worst performing trial was weighted most significantly ($\times 1$) and its best performing trial was weighted least significantly ($\times \frac{1}{20}$). This weighted sum was then normalised such that the maximum possible fitness was 1 and the minimum was zero.

Selection of parents was fitness rank proportionate, using Baker's linear ranking method (with maximum expected offspring equal to 1.1) and Baker's stochastic universal sampling (Baker, 1987). Offspring were generated asexually. The offspring genome was mutated by applying additive Gaussian perturbation to the value of each gene. A vector of 18 independent samples from a standard Gaussian distribution was normalised such that it summed to a mutation magnitude value drawn from a Gaussian distribution with zero mean and variance equal to 0.2. This vector was then added to the vector of offspring gene values. Any mutated gene values lying outside the range $[-1, +1]$ were clipped to the nearest legal value. If the new offspring did not achieve a fitness greater than that of its parent, its place in the new generation was taken by its parent.

For convenience, values used for each parameter described above are shown in Table 2. To the best of our knowledge these are equivalent to those employed by Campos and Froese (2017).

For the two-way communication task, the score for each individual trial was calculated on the basis of the distance between the true target location and the final positions of both agents as follows:

$$S_i^T = \max(0, 1 - |A_i^T - G^T| - |B_i^T - G^T|) \quad (4)$$

Where S_i^T is the score achieved on trial T by individual i from the current population of networks, A_i^T and B_i^T are the locations of two simulated agents controlled by network i (A_i and B_i) at the final time step of trial T , and G^T is the location of the true target during trial T . This fitness function requires both agents to navigate to the true target location,

Table 2. Parameters for the One-Way Communication Task Scenario, CTRNN and Genetic Algorithm.

Population size	: 50
Generations of evolution	: 5000
Genome length	: 18 real values
Legal gene allele range	: [- 1, +1]
CTRNN size	: 3 neurons
Tau (τ) range	: [50, 100]
Weight range	: [- 16, +16]
Bias (θ) range	: [- 16, +16]
Max. expected offspring	: 1.1
Mutation magnitude	: 0.2
Trials/fitness assessment	: 20
Trial duration	: 300 units of time
Integration step	: 1 unit of time
Contact sensor range	: 0.4 distance units
Other sensor ranges	: Unlimited
Motor output	: Linear mapping
Maximum speed	: 0.01 distance units/time unit

giving higher scores when both agents perform well. This means that if the agents are unable to reach a consensus regarding which target is best then they will always receive a low score and therefore joint coordination of their behaviour is essential.

For each individual network, fitness was calculated on the basis of performance across 20 trials, 10 featuring a true target with a random *positive* location, and 10 featuring a true target with a random *negative* location. For each generation, the same 20 trials were used to evaluate each network in the population. The fitness of an individual network was calculated as a weighted sum of the scores that it achieved across *either* the 10 trials featuring a true target with a positive location *or* the 10 trials featuring a true target with a negative location, whichever was worst. Again, each score’s weight was equal to the reciprocal of its rank (ascending) within the list of 10 trial scores achieved by that individual on the chosen side of the environment, thus, ensuring that the individual was assessed on its worst performing side of the environment and that its worst performing trial on that side was weighted most significantly ($\times 1$) and its best performing trial on that side was weighted least significantly ($\times \frac{1}{10}$). Again, this weighted sum was normalised such that maximum possible fitness was 1 and the minimum was zero.

Results

One-Way Referential Communication

Figure 5 shows the highest fitness value in the population across five representative 5000-generation evolutionary runs. One commonly evolved mediocre, degenerate solution to the task is for the receiver to ignore the signaller and

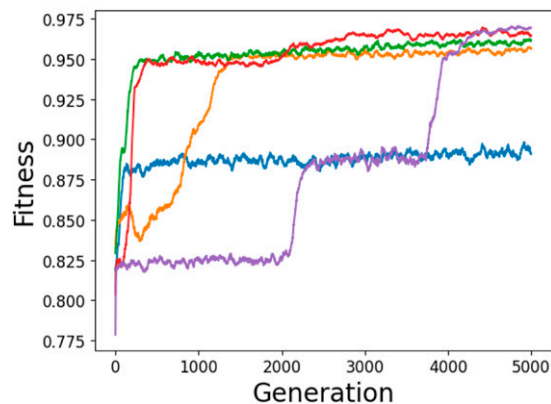


Figure 5. Maximum population fitness over time for five representative evolutionary runs of the one-way communication task, smoothed by taking a 50-generation rolling average.

move to the average target location. This achieves an average fitness of 0.816 and corresponds to the behaviour represented by the purple line between generation 0 and 2000. There also appears to be a significant fitness plateau at around 0.875, which some runs struggle to surpass (e.g. the one represented in blue). Because of the increased weighting given by the fitness function to the lowest scoring trials, in order to gain a higher fitness score the network must be able to generalise its performance across the range of target locations. Manual analysis of the population represented in blue revealed networks that can only successfully deal with a limited range of target locations.

Most evolutionary runs were able to achieve solutions with a fitness $\cong 0.95$, which corresponds to recognisably competent signalling behaviour. Here, we will analyse two of the best performing evolved solutions. Network 1 (Figure 6) and Network 2 (not shown) enable the signal receiver to finish trials within 0.05 distance units of the target location in 99.2% and 98% of cases, respectively. By comparison, the best performing network reported by Campos and Froese (2017) managed to achieve this degree of accuracy in 97% of the trials.

Figure 7 demonstrates the behaviour of Network 1 on three distinct trials, each starting from the same initial conditions but confronted with a different target position. The figure shows the locations of each agent and the target, and also the values for each of the agents’ sensors throughout the trials. The strategy performed by Network 1 has two phases. Firstly, for roughly 100 units of time both agents move in a positive direction. As a consequence of the fact the signaller is prevented from moving outside its ‘signalling zone’, this enables the two agents to separate and extinguishes differences caused by their initial random starting positions. Subsequently, both agents reverse direction and move back towards the origin before reversing

their direction of motion a second time and moving towards the target. The signaller indicates the location of the target by performing this reversal at a time related to the magnitude of the target location co-ordinate, thereby altering the

moment at which the two agents regain contact with each other. This allows the receiver to modulate the point at which it makes its second reversal of motion such that it is able to finish the trial at the target location.

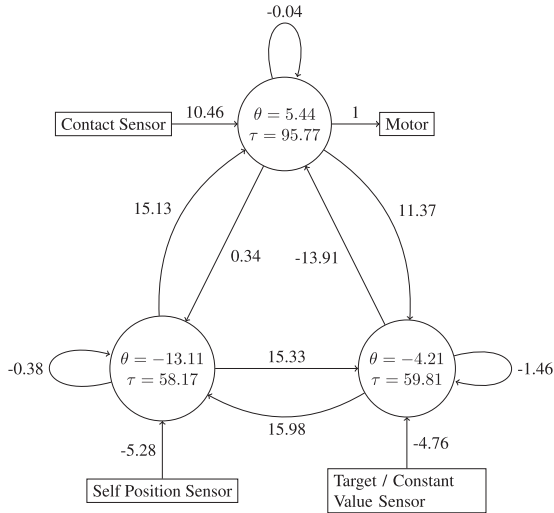


Figure 6. The weights, biases and time constants of Network 1, a high fitness individual for the original one-way communication task.

An alternative evolved strategy (Network 2) adopts a similar scheme, but exploits the fact that only the receiver is capable of moving to negative locations. Network 2’s behaviour also features two phases. Firstly, the signaller and receiver travel in a negative direction. The signaller is prevented from travelling to negative coordinates (which are outside the signalling zone) allowing the receiver to move out of contact sensor range, which enables the agents to extinguish the effect of variation in their initial conditions caused by their random initial starting positions. Both agents then reverse direction, moving positively towards the target (although the signaller cannot move further than the upper boundary of the signalling zone). The signaller reverses direction a second time at a moment dictated by its sensory reading of the target location. This serves to determine the amount of time for which the two agents are within contact range of each other, thereby enabling the receiver to modulate its outward motion such that it reaches the target location at the end of the trial.

Figure 8 depicts the overall accuracy of the Network 1 and Network 2 strategies. The left-most pair of graphs show the mean distance between the receiver and the target

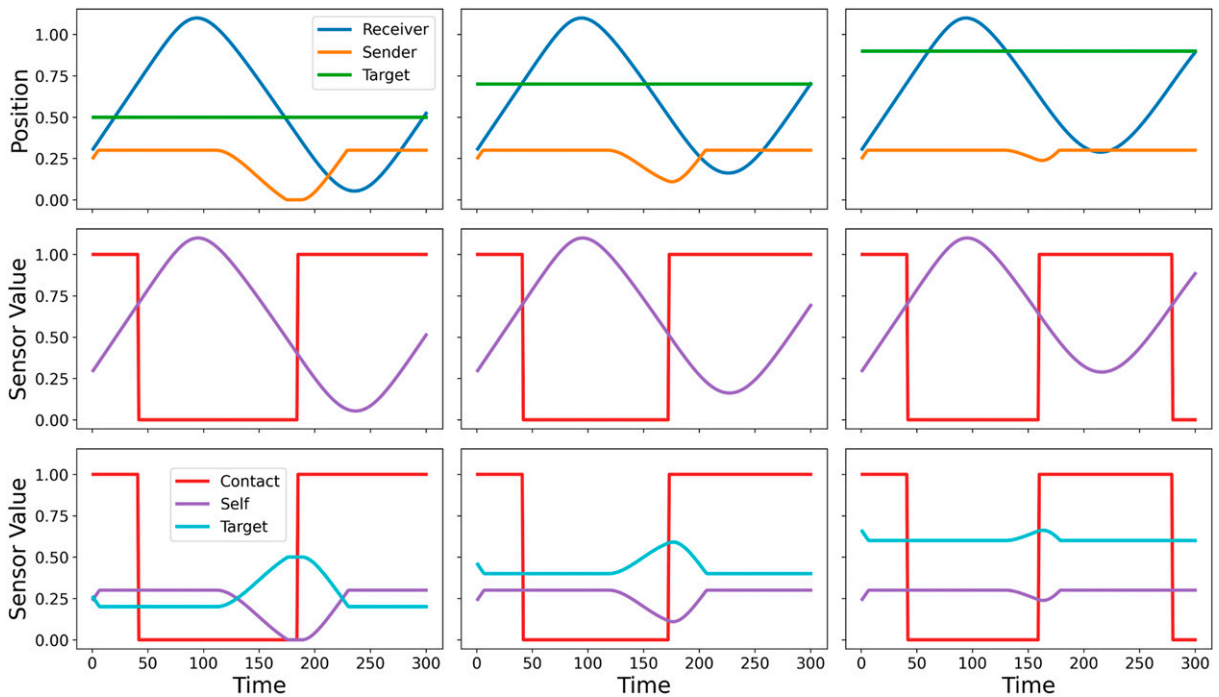


Figure 7. The behaviour of Network 1 on three example one-way communication trials (one per column) that differ only in terms of the target location. The top row presents the positions of the agents and the target throughout the trial. The middle and bottom rows present the manner in which sensor values vary over the course of each trial for the receiver and signaller, respectively.

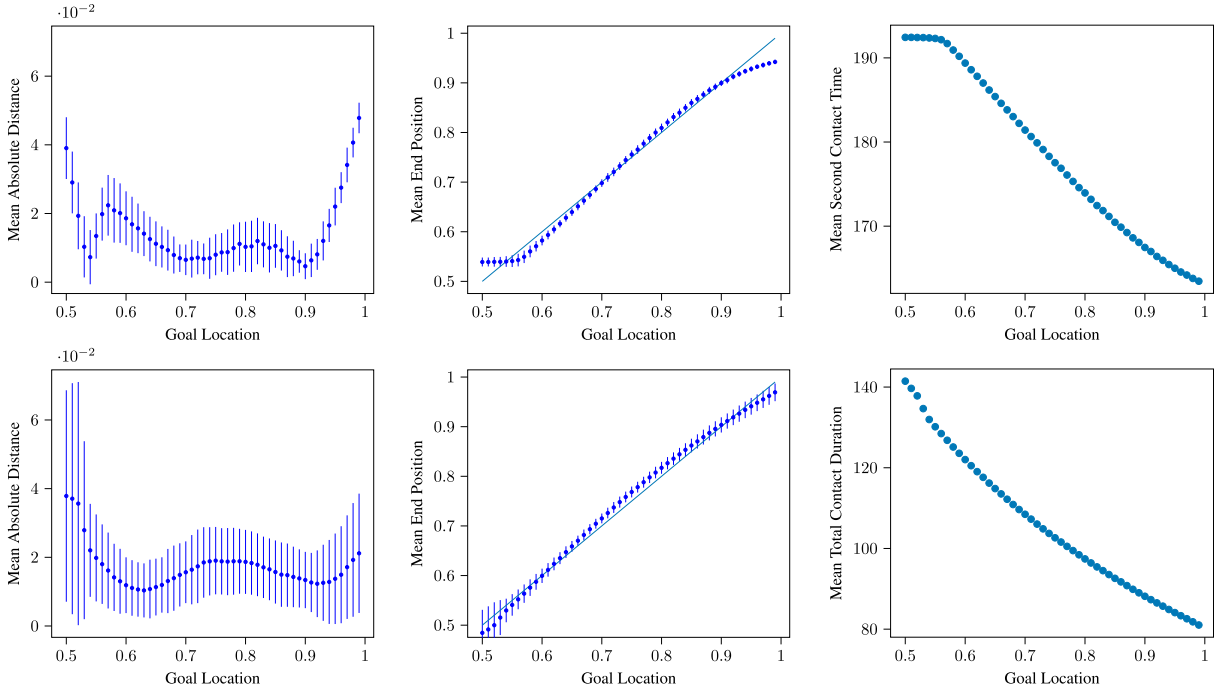


Figure 8. *Left:* The mean absolute distance between receiver and target location at the end of a trial, plotted against the location of the target during the trial for Network 1 (top) and Network 2 (bottom). Error bars show one standard deviation. *Centre:* The mean receiver position at the end of a trial, plotted against the location of the target during the trial for Network 1 (top) and Network 2 (bottom). Error bars show one standard deviation and the line $y = x$ depicts perfect performance. *Right:* For Network 1 (top), the average time at which contact is re-established after the first period of separation for signaller and receiver, plotted against the location of the target. The Pearson product-moment correlation coefficient, $r = -0.9959$. For Network 2 (bottom), the total amount of time during which the agents are within contact sensor range, plotted against the location of the target. The Pearson product-moment correlation coefficient, $r = -0.9892$. Each data point summarises 100 trials sampling all signaller and receiver initial locations drawn from the set $\{0, 0.03, 0.06, 0.09, 0.12, 0.15, 0.18, 0.21, 0.24, 0.27\}$ (with replacement). (Note that the vertical scales differ between plots.)

location at the end of a range of trials for Network 1 agents (top) and Network 2 agents (bottom). For the vast majority of the target range, this distance is well below the arbitrary success threshold of 0.05 employed by Campos and Froese (2017). The worst performance tends to be associated with the most extreme target locations. The central pair of graphs shows the mean final position of the receiver over a large range of initial conditions for Network 1 (top) and Network 2 (bottom). The receiver position closely approximates the target location with some slight divergence at the extremes of the target location range. The standard deviation around the performance is small, indicating that the Networks are able to achieve a high performance consistently, regardless of variation in the agents' initial starting locations. The right-most pair of graphs support the hypothesis that the precise timing of contact being re-established between the two agents (Network 1) and the precise duration of sensory contact between the two agents (Network 2) is likely to be strongly involved in achieving successful task behaviour since the correlation between these aspects of the agents' joint behaviour and the true location of the target are extremely strong.

In order to further understand the behaviour of the two networks, the neural states of agents controlled by Networks 1 and 2 have been plotted for a range of trials featuring targets at one of six different locations (Figure 9). Note that unlike the behaviour of the signaller, which gradually diverges in a manner that depends on the target location, the behaviour of the receiver is identical across these trials until a certain point in time. This corresponds to the initial 'separation phase', during which the behaviour of the signaller has no impact on receiver movement. After this point however, the behaviour of the network diverges, with the neural states branching in a way that depends on the signaller's location in a way that strongly correlates with the target location. It may be tempting to identify the sharp discontinuities in the neural trajectories as 'decision points' corresponding to moments at which the agents achieve communication. However, these discontinuities correspond to points at which either (i) the signaller reaches a signalling zone boundary, (ii) the signaller and receiver separate to the extent that they lose contact with each other. Removing the signaller

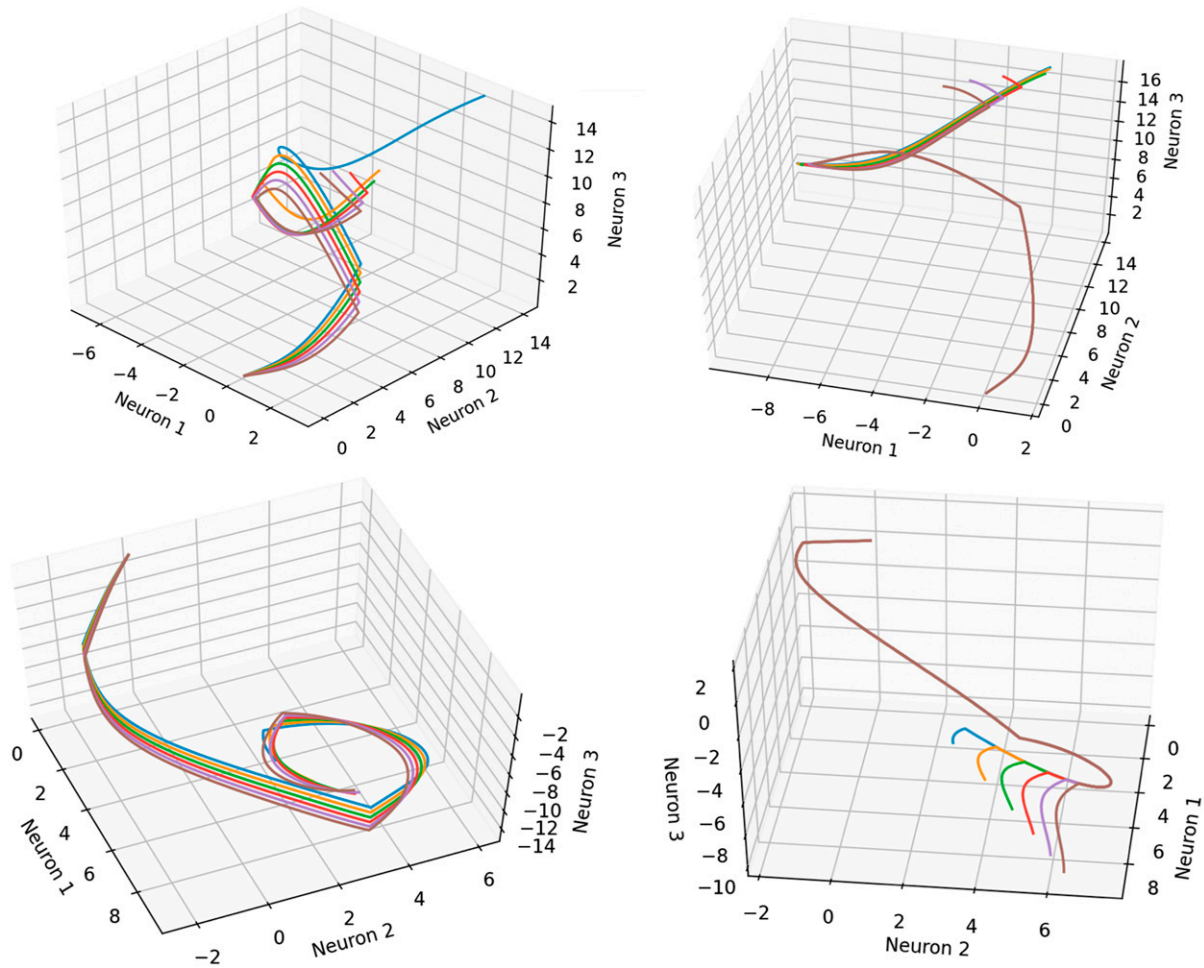


Figure 9. The neural dynamics for the signaller (left) and receiver (right) controlled by Network 1 (top) and Network 2 (bottom), depicted for six independent trials of the one-way communication task with target locations ranging from 0.5 (blue) to 1.0 (brown) in steps of 0.1.

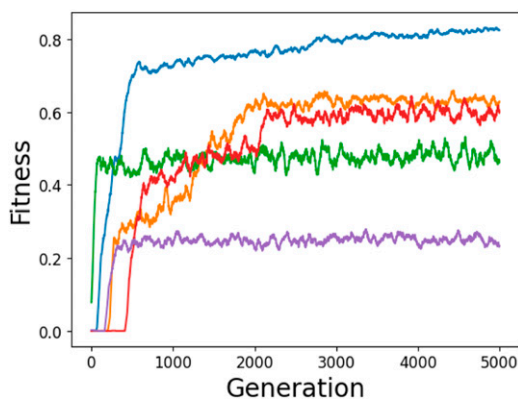


Figure 10. Maximum population fitness over time for five representative evolutionary runs of the two-way communication task, smoothed by taking a 50-generation rolling average.

from a trial at any point prior to the agents losing contact for the final time tends to have a (negative) impact on receiver behaviour. Thus, rather than being associated with a discrete event, agent communication is a temporally extended, continuous interaction reliant on the timing of the onset of contact and/or the duration of this contact.

Two-way Referential Communication

Solutions for the two-way communication problem were significantly harder to evolve than for the original one-way communication task, as shown by Figure 10. Not only do the evolutionary runs for the two-way communication task tend to exhibit larger variance in fitness across consecutive generations, but there is also a much larger variance in the range of fitness values achieved overall. For the original

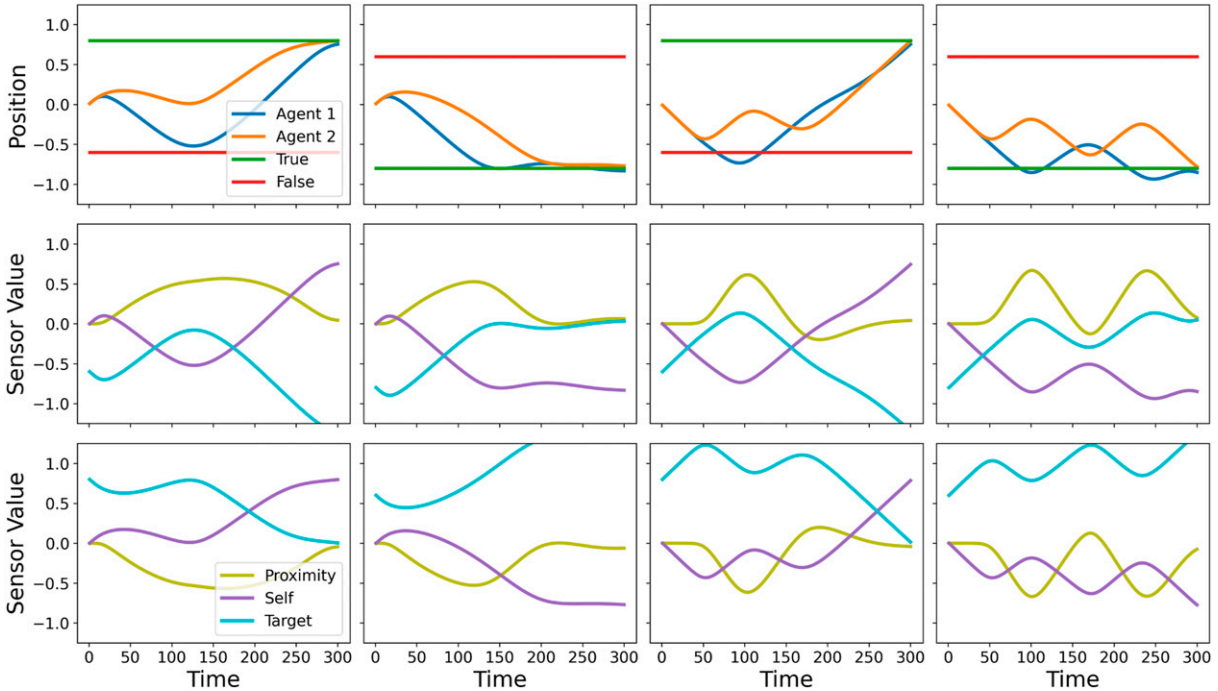


Figure 11. The behaviour of two successfully evolved CTRNNs on two examples of the two-way communication task. The top row presents the positions of the two agents and the two targets (the true target and the false target) throughout the trial. The middle and bottom rows present the manner in which sensor values vary over the course of each trial for each of the two agents, respectively. The left-hand pair of columns show Network 3’s behaviour on two representative trials, one in which the true target is in a *positive* location and one in which the true target is in a *negative* location. The right-hand pair of columns show Network 4’s behaviour, again on a trial with a positive true target and a negative true target, respectively.

one-way communication task, a non-zero fitness will always be awarded to an agent that stays still or moves less than 150 distance units in the positive direction. However, for the two-way communication task most evolutionary runs feature several initial generations in which every genome achieves a fitness of zero.

As the fitness function is based on the combined distance of both agents from the true target, if the agents travel to different sides of the environment their fitness will almost always be zero. Similarly, if the agents always go to one side, regardless of whether the true target is on that side or not, then the overall fitness will once again be zero, as the fitness function only considers performance on the agents’ worst performing side. Lastly, if the agents do not move at all, they will again receive zero fitness as the combined distance to the true target will be greater than 1. This means that a randomly initialised population is far less likely to include strategies that achieve non-zero fitness, and that the first few generations will therefore often perform very poorly. Consequently, once a mutant strategy with non-zero fitness arises, it is likely that it will reproduce rapidly leading to a strongly converged population, despite the relatively weak selection pressure that is implemented. Nevertheless, although the median best

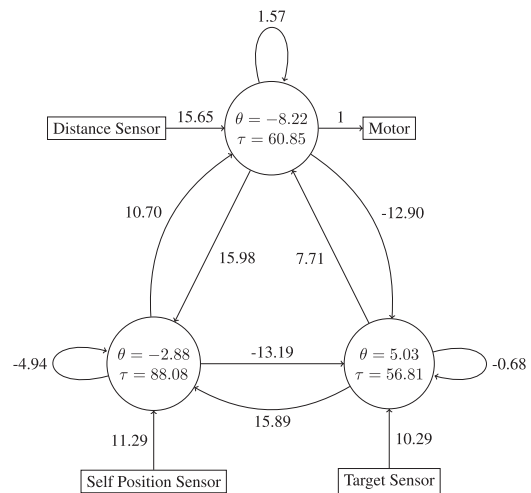


Figure 12. The weights, biases and time constants of a high fitness individual for the two-way communication problem (Network 4).

fitness was close to 0.6 across all 20 simulation runs, some networks managed to achieve higher values, up to 0.9 which corresponds to recognisably communicative performance.

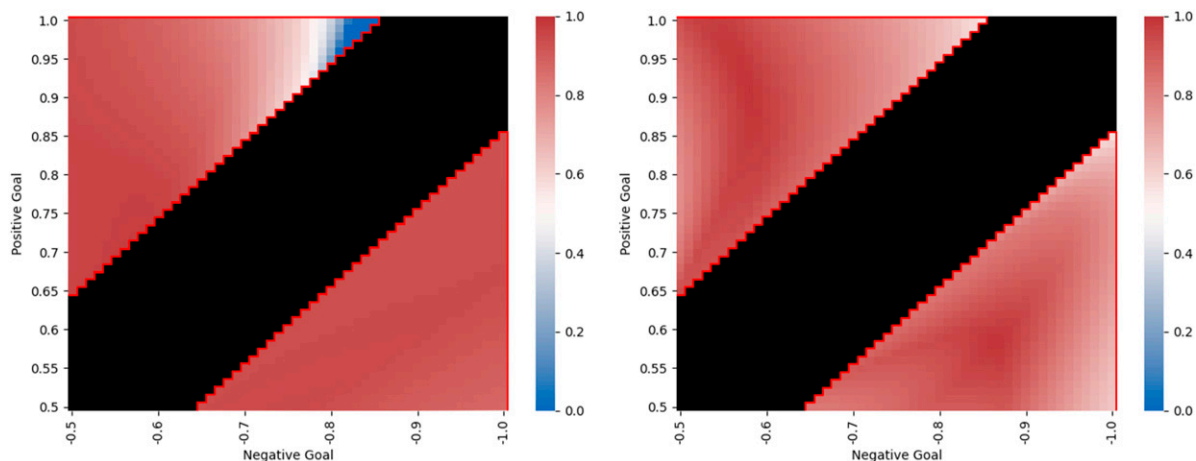


Figure 13. Heat maps showing variation in two-way communication performance across a range of trials for Network 3 (left) and Network 4 (right). The triangular regions enclosed by a red line correspond to trials experienced during evolution of Networks 3 and 4. The central black region corresponds to trials that the evolving populations were not exposed to because the difference between the absolute value of the two target locations is less than 0.15.

The behaviour of two successful networks is presented in Figure 11. Due to the continuous range of values that the proximity sensor can take (by comparison with the binary contact sensor used for the one-way communication task) the two-way communication behaviour has the potential to be far more subtle than that observed in the previous task. As a consequence, the behaviour of the evolved networks can be significantly harder to analyse. However, as with the original one-way communication task, the evolved behaviour typically appears to have two phases. In the first phase, movements are performed that allow each agent to share the location of their own goal with the other agent (the ‘communication phase’). In the second phase, both agents move towards the true target (the ‘decision phase’).

For Network 3, it appears that the communication phase culminates at around $t = 120$ when Agent 1 (which is able to sense the location of the positive target but does not initially know whether this is the true target or not) decides to either alter its trajectory by moving towards the positive target (causing Agent 2 to accompany it) or, alternatively, decides to continue to follow Agent 2 towards the negative target. The former outcome obtains in situations where Agent 1’s co-ordinate at the $t \approx 120$ ‘decision point’ is higher than some threshold, but the precise mechanism by which the agents are able to make this ‘decision’ correctly across the range of trials that they must deal with is not obvious from observing their external behaviour. The behaviour of Figure 12 is significantly different. The agents commence each trial by moving towards the negative target in a sinusoidal pattern that relies on them making use of their continuously varying proximity sensors. However, if the co-ordinate of the third turning point of this sinusoidal behaviour is greater than the co-ordinate of their first turning

point (which occurs at around $t = 50$ in almost every trial), then the agents cease this sinusoidal behaviour and travel directly towards the positive target. Again, whilst it is possible to correlate a feature of the agents’ joint trajectory with the ‘decision’ to choose the positive target versus the negative target, the extent to which this account captures the true causal mechanisms in play is not clear.

Figure 13 shows the performance of Networks 3 and 4 across a range of trials in which the positive and negative target locations vary between $[0.5, 1]$ and $[-0.5, -1.0]$, respectively, exhaustively sampled at intervals of 0.01 units. Scenarios that were invalid during evolution due to the absolute values of the target locations being too close together are coloured black. Both heat maps show good performance across the entire range of scenarios. For Network 3, the lowest scoring scenarios are in a small region near $(1.0, -0.85)$ where performance falls to zero in the worst cases. When the true target is negative, fitness scores are in general higher, indicating that the agents are using a more efficient method of communication for this half of the problem. By contrast, Network 4 is capable of stronger performance overall. Notably, there are no scenarios for which the fitness score falls below 0.56, meaning there is strong performance across the whole range of scenarios to which the networks’ lineage was exposed during evolution.

In order to determine how well these networks generalise, Figure 14 plots their two-way communication performance across an extended range of scenarios that involve target locations outside the range employed during evolution, and also removes the lower bound on target separation (i.e. the difference between the absolute values of target locations is no longer constrained to be greater than 0.15).

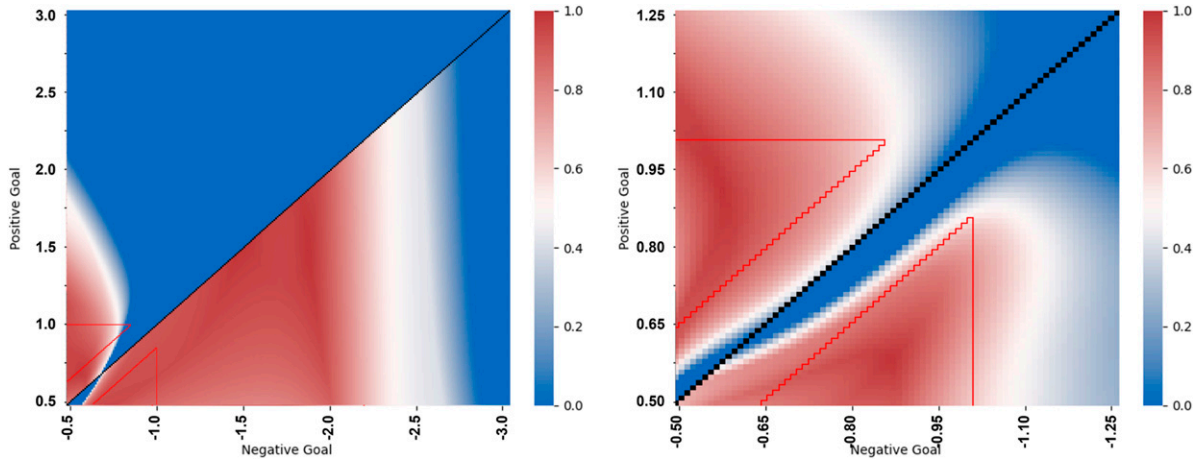


Figure 14. Heat maps showing the extent to which Network 3 (left) and Network 4 (right) are able to generalise their two-way communication performance beyond the scenarios to which their ancestors were exposed during their evolutionary history. Again, the triangular regions enclosed by a red line correspond to the trials employed throughout the evolutionary runs. Note the scales on each heat map are different and also differ from those employed in Figure 13.

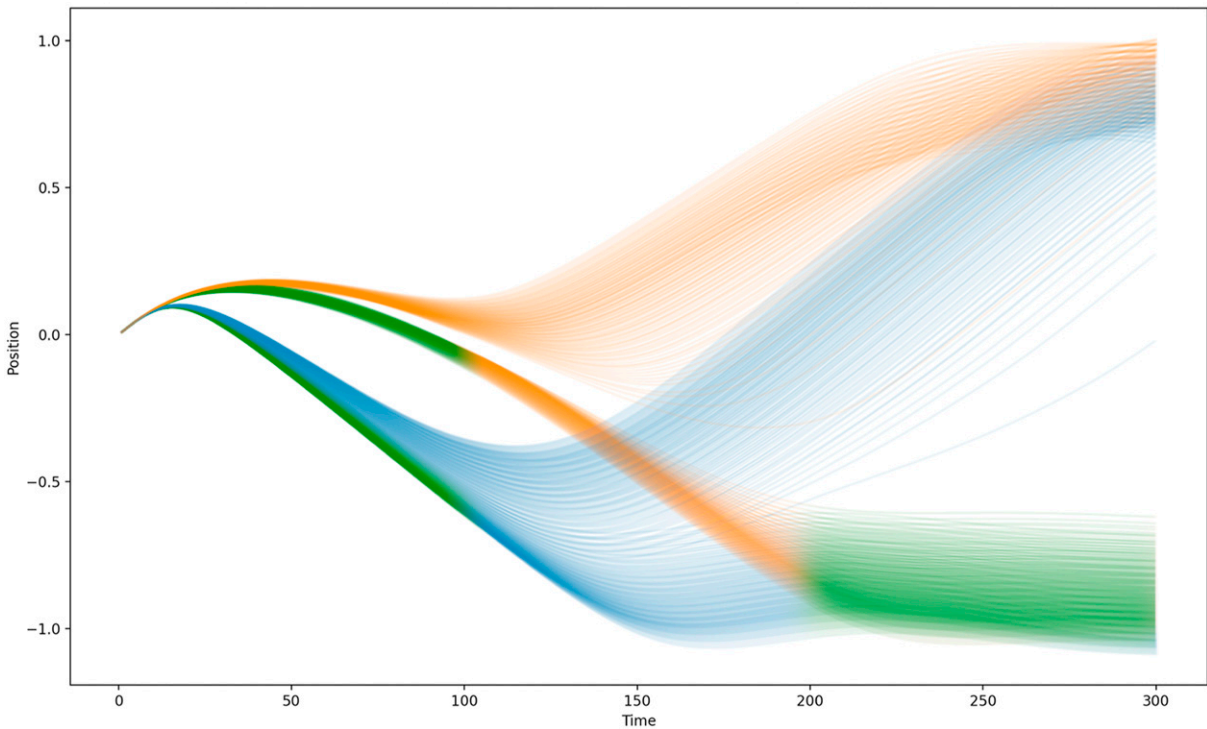


Figure 15. Trajectories for Agent 1 (orange) and Agent 2 (blue) using Network 3 to solve the two-way communication task across a range of trials with different positive and negative target locations. Tasks for which the true target was in the negative target range are coloured green for the first and last 100 time steps.

Scenarios that were employed during evolution are surrounded by a red line and scenarios for which targets are an equal distance from the origin are coloured black (as correct behaviour is undefined in this case). Whilst Network 3 can

cope with negative targets that are far more extreme than those experienced during evolution (which is consistent with its higher performance for negative targets in general), overall, both heat maps indicate that the evolved solutions

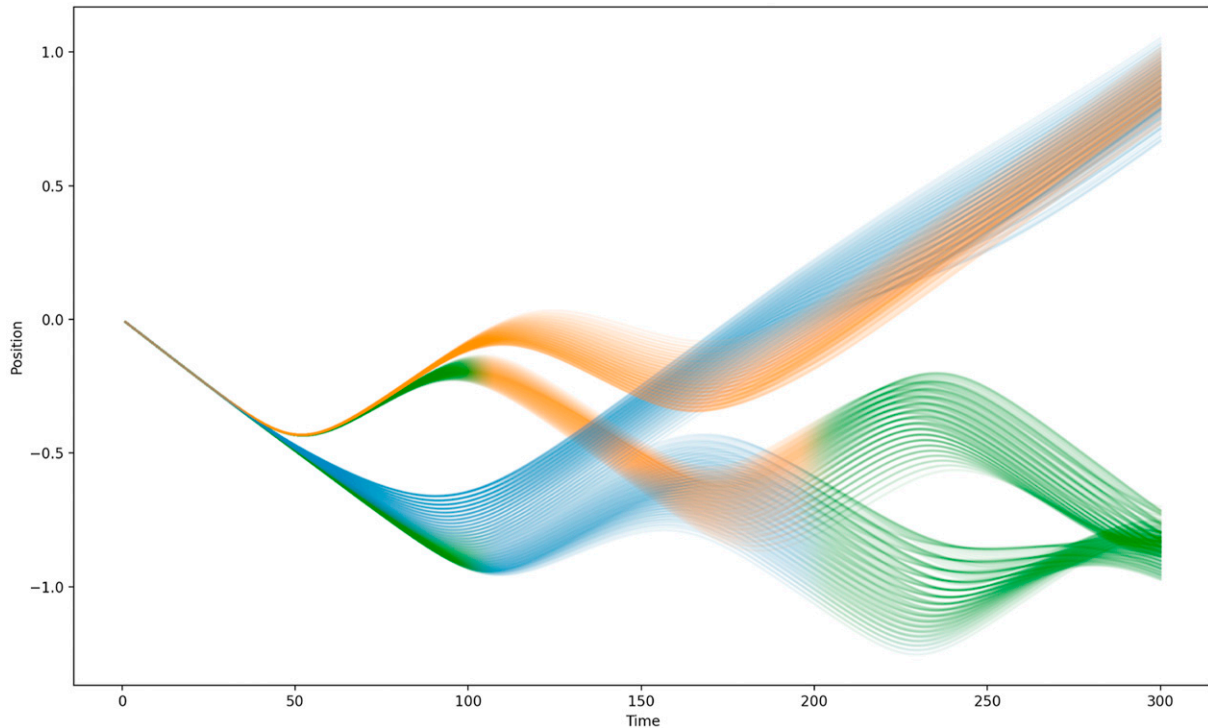


Figure 16. Trajectories for Agent 1 (orange) and Agent 2 (blue) using Network 4 to solve the two-way communication task across a range of trials with different positive and negative target locations. Tasks for which the true target was in the negative target range are coloured green for the first and last 100 time steps.

can generalise only to some limited and unpredictable extent, with performance tending to decline steeply towards zero for increasingly novel scenarios.

In Figure 14, a blue region of low fitness can represent one of two different failure modes. When performance changes abruptly from high fitness to low fitness (i.e. from red to blue), this can represent agents persisting with visiting one side of the environment when the target is now on the other side, i.e., consistently making the wrong choice regarding which side to navigate towards. Alternatively, when fitness gradually degrades over a region of the heat map (i.e. from red through white to blue), this represents agents becoming increasingly unable to accurately reach the target location by the end of the trial, but still tending to predict correctly on which side of the environment the target location lies. This may be caused either by a failure in the communication scheme’s ability to direct agents to targets beyond a certain point, or by the fact that limits on agent speed prevent them from having enough time to reach the target location before the end of a trial, despite in principal having the ability to travel to it accurately.

Earlier it was presumed that the agents first signal their private information about a target location and then decide

which target they should travel towards based on the influence that this signalled information has had on their trajectory. Closer examination of the agent behaviour demonstrates that this discretised ‘signal-decide-act’ interpretation is not correct.

Figures 15 and 16 depict, for Networks 3 and 4, respectively, the trajectories of both signallers and receivers across every pair of target locations in the range $[0.5, 1]$ for the positive goal and $[-0.5, -1]$ for the negative goal sampled at intervals of 0.01. (Note that this set of trials includes ones in which the magnitude of the true target’s co-ordinate lies in the range $[0.51, 0.65]$, i.e. trials that would not be encountered by the agents during evolution.) Trials for which the negative target was in fact the true target have been shaded green for the first and last 100 time steps. These figures show that the behaviour of both agents is determined by the true target location far earlier than described previously. For both Network 3 and Network 4, the behaviour of both agents is already influenced by whether the true target is positive or negative by $t \approx 50$, implying that the agents have already begun to ‘decide’ which side to navigate towards before this point.

The method by which the agents exchange information exploits the fact that they are initialised in the same location

and in the same the neural state and that their proximity sensors are both noiseless and precise. This means that even subtle differences in the relative trajectories of the agents (that reflect the absolute value that they receive from their target distance sensor) can be exploited by the agents to effectively assign a ‘leader’ role to the agent with direct knowledge of the true target location and a ‘follower’ role to the other. Whilst there remains a considerable amount of variation in the trajectories subsequent to this very early differentiation of agent roles, this can be accounted for in terms of (i) the leader Agent following a target-location-specific trajectory that converges on the directly sensed target location at the end of the trial, and (ii) the follower Agent timing its convergence on the location of the leader such that they also coincide at end of the trial. Again, removal of either agent during the trial does damage the ability of the other to complete the task successfully, so there remains a significant mutual coupling between the agents which is not fully captured by the leader/follower labels.

Perhaps somewhat remarkably, then, despite the effort made in the literature and the current study to introduce a series of increasingly complex signalling tasks intended to demonstrate increasingly sophisticated displaced referential communication, we have arrived at successful evolved solutions that still strongly resemble the leader/follower co-ordination described in Quinn’s (2001) original model of evolving communication without dedicated communication channels. Whilst that model did not include an explicit target about which the agents needed to signal, in other respects the behavioural interpretation of the evolved solutions are remarkably similar, each involving an interplay of proximity-based oscillation and following behaviours.

Discussion

In this section we will first consider the status of the evolved communication demonstrated in this paper before discussing the future prospects for this style of work to shed light on the minimal cognitive basis for communication as complex as human language.

The bidirectional referential communication evolved by the agents in this study is effective in allowing them to solve the evolutionary task that they were confronted with. This is a positive result given the aims of the study. However, the signalling system evolved to achieve this success is limited in several respects.

Firstly, the signalling behaviour is fundamentally *analogue* or continuous, in the sense that the signal semantics amount to a mapping between some scalar property of joint agent behaviour (e.g. the amount of time that the two agents spend within sensor range of each other) and some scalar property of the environment (i.e. the co-ordinate of the target). Secondly, the signalling is also *holistic* in the sense that this mapping lacks structural complexity, being

essentially one-dimensional and monotonic. Thirdly, the signalling behaviour has only a limited degree of *productivity*. Whilst the location of a target lying within the normal range experienced by the agents during their evolution can be successfully communicated even if this precise location has never previously been experienced or communicated during evolution, there is no ability to signal about an open-ended set of referents. Fourthly, the signalling exhibits only a restricted kind of *referent displacement*. Whilst a target that is not present to the senses of one agent can be communicated about successfully by the other, neither temporal nor counterfactual displacement are demonstrated.

Finally, like much animal communication, for the evolved agents there is no effective separation between an *imperative* interpretation of signal meaning (‘Do X!’) and an *indicative* interpretation of signal meaning (‘Believe Y!’). For example there is no in principle way of attributing the meaning ‘The target is at location x!’ to an agent’s signalling behaviour without also attributing the meaning ‘Go to location x!’ – these meanings are coupled to the extent that they are inseparable (Millikan, 1984).

Whilst the honey bee dance language shares each of these four limitations to some degree, human language is different. It is not limited to analogue, holistic signals, instead employing utterances that comprise *discrete* parts (words) that contribute to the meaning of an utterance in a way that depends on their *configuration* within the utterance (i.e. meaning derives from a compositional grammar of some kind). Whilst the signalling system evolved here does generalise across a range of target locations, the recursive, particulate, compositional nature of human language means it is productive in a much more profound sense, having limitless potential to say truly novel things. The referent of a human language utterance may also be displaced arbitrarily with respect to the speaker and listener: spatially, temporally and even counterfactually as when an imaginary subject is discussed. Finally, human language also has the capacity to decouple imperative and indicative moods, specifically being able to communicate particular information without simultaneously conveying an attendant instruction to act immediately in a particular way: for example ‘Every triangle has three sides’, ‘My name is John’.

These distinctions between human language and the agent communication evolved here will be returned to below when we consider potential future work. Beforehand, what can we infer about the cognitive basis for communication from the results presented here (and the results of the previous studies that the current study builds upon)?

The results presented here demonstrate that successful bidirectional referential communication is possible for agents that are *extremely simple*. A network of only three idealised model CTRNN neurons is sufficient for an agent to operate as both signaller and receiver in a bidirectional communication task that requires two pieces of information

to be communicated, compared and acted on. The evolved agents are able to solve the task by exploiting a ‘signalling channel’ that they construct from their physical movements as detected by their simple distance sensors.

It is tempting to describe this evolved communication behaviour in terms of a series of discrete stages that mirror a traditional view of the psychology of communication: an agent first signals its own private information and receives the signal sent by its partner, then compares the two pieces of information, decides whether to lead or follow their partner, and then carries out the chosen behaviour. However, a closer analysis of the behaviour (reported in the previous section) reveals that there is little basis for projecting this sequence of stages onto the agents’ joint activity which more closely resembles a continuous period of mutual sensory-motor modulation achieved by carefully calibrated structural coupling between the two agents.

This kind of dynamical systems interpretation of the evolved communication behaviour allows little room for explanations couched in terms of ‘representations’ or ‘symbols’. This is true in two distinct senses.

Firstly, there is no real *need* for this kind of explanation. A sufficiently detailed dynamical systems analysis is in principle capable of providing a full and complete causal-mechanistic account for the evolved behaviour without any need to employ explanatory entities that are representational or symbolic *for the agent*. Such entities would at best supervene on a full dynamical systems account. Secondly, there is no real *purchase* for this kind of explanation. Whilst we might point to events within the simulation that have representational content for ourselves as observers, it is not clear that there exist internal, external or joint behavioural phenomena that could correspond to representations or symbols for the agents themselves, denying these type of explanatory entities even a role in some supervenient explanation intended to be layered on top of a full dynamical systems account.

Taken as a whole, then, this picture usefully draws attention to the fact that a quasi-linguistic interpretation of cognitive innards is not necessary or automatically appropriate (Van Gelder & Port, 1995), and that choosing to use ‘internal representations’ as part of a causal explanation of cognitive behaviour is an explanatory strategy that needs to be carefully considered and justified. Agents that are sophisticated enough to demonstrate the cognitive behaviours that we are interested in may nevertheless not be ‘sophisticated enough’ to warrant explanations of this kind.

However, this way of framing the results also leads to a problematic tension between the studies undertaken so far and their presumed ultimate explanatory target: the (minimal) cognitive basis for communication that is as advanced as human language. Human communication involves the production and consumption of truly symbolic representations. The evolved communication reported in the studies

considered here does not. Can developing such studies further ever shed light on human language use? Presumably there will always be a dynamical systems theoretic account of any evolved communication, and this account will undercut talk of symbols, concepts and representations in the same way that a physical account of the brain that is articulated in terms of atomic collisions would make no mention of beliefs and desires. However, in order for the research programme being pursued here to shed light on human cognition, it must be the case that bridging explanations that link dynamical systems accounts to cognitive phenomena are *possible*, at least in principle (Bullock, 2004), and it must be the case that the agent frameworks and simulations being explored here are capable of scaling up to generate behaviour that is sophisticated enough to demand such bridging explanations, despite them not being models of particular real-world creatures or particular real-world brains (Bullock, 2009).

Versions of this ‘scaling-up’ tension are common across bottom-up, bio-inspired, ‘nouveau’ AI research. For example the insect-inspired robotics field, which sought to build towards robots capable of human-level sophistication by starting from very simple robots inspired by ants and hover flies, was vulnerable to similar criticism amounting to: you would not reach the moon by climbing successively taller trees (Matarić & Cliff, 1996). Are the evolved agents reported here just climbing a slightly taller tree than previous studies? Is human language an impossibly distant moon? With these questions in mind, it is worth asking to what extent could further work of the kind reported here deliver more sophisticated communication that resembles full-blown human language? What barriers must be overcome?

Millikan (1989a) provides a useful list of representational properties that distinguish true language from more simple animal signalling. They connect directly with the short-comings of the evolved agent signalling reported here, mentioned at the start of this section, and together provide a kind of grand challenge for research into the evolution of agent signalling.

Self-Representing Elements

Simple animal signals map onto their referent in a simple way, often relying on some property of the signal to represent the self-same property of the signal’s referent. For instance, an alarm call might use a high pitched shriek to indicate that a predator is approaching. Implicitly, the alarm call also indicates the *where* and the *when* of the approaching predator (“right here” and “right now”, respectively) by virtue of the *where* and the *when* of the signal also being “right here” and “right now”. Such a system is therefore unable to represent a referent with a where and when other

than “now” and “here”. By contrast, human language is able to convey “We are being attacked tomorrow” or “We are being attacked in Paris”.

Storing Representations

If signals that map onto a *where* and a *when* other than the here and now are to be useful, then their meaning must typically somehow be persistent within the agent such that the agent can choose to act on the meaning when the time and place are right. By contrast, many simple animal signals are ephemeral, being relevant only instantaneously and therefore not requiring storage. The bee dance language is a departure in that the referent of the signal is spatially displaced from the signaller and observer. The meaning of a bee dance must continue to influence the observer once it has flown away from the dance floor. Similarly, the evolved agents reported here must often exhibit behavioural persistence in reaching a location that is outside their signalling zone. However, for both honey bees and our evolved agents, the *when* of a signal’s referent is always *now* which means that it is premature to talk of storing representations of target locations. It is sufficient to talk in terms of heading off immediately in a signalled direction at a signalled intensity. Our evolved agents need not store a representation of their target location that must be consulted as they move towards it. Rather they need only set off at the right speed, such that the end of the trial coincides with them reaching the correct location.

Indicative and Imperative Representations

As mentioned already, language utterances may have a mood that is either imperative or indicative or a mixture of both. This flexibility is not true of much animal signalling where the meaning of a signal is both indicative “Nectar is at location X” and imperative “Go to location X”. That the cognitive story for such cases involves no information processing is therefore hardly surprising. There is little cognitive work to be done on the signal. It maps directly to an imputation to act. A cognitive story with little room for information processing has limited justification for being called “representation hungry”.

Inference

When the act of receiving and understanding a signal can be decoupled from taking immediate external action, the door is opened for more sophisticated internal cognitive behaviour. If a signal’s indicative mood is decoupled from its imperative mood, some kind of inference will be required in order to recouple its meaning with downstream action. Here then perhaps, to the extent that the contexts for this downstream action are

many and varied, is where the real hunger for something like internal representation may lie: a cognitive ability to generate intentions (and actions) from beliefs.

Acts of Identifying

Once a signal’s meaning is decoupled from the where and when of its immediate sensory-motor context, and is to be retained for unspecified usage in some other where and when as instigated by inferential processes, the challenge arises of re-identifying its referent both in the world and internally. This is not a challenge for our evolved agents which simply set off immediately with a direction and speed that ensures they arrive where they need to arrive when they need to arrive without needing to recognise that they have so arrived. By contrast, acting on a signal such as “I will mark the traitor’s door with red just after midnight” requires much more of the relationship between the internal state and the downstream behaviours that it enables. The referent as indicated somehow in the trace of an indicative signal must somehow be identified with the referent as identified in the sensory world of the agent. Moreover, the potential for the referent of multiple different imperative or indicative signals to be identified as one and the same underpins much sensible behaviour (as when it transpires that “Sam’s door” is also “The traitor’s door”). Achieving this “substitution of similars” (what Stanley Jevons called “a dark and inexplicable gift”, Bullock, 2008) requires at minimum that internal states be articulated as part of some system, rather than merely mapping holistically onto the world.

Negation and Propositional Content

Finally, Millikan argues that natural signalling systems, such as the bee dance language, lack *negation*. When two bee dances indicate two competing locations it may well simply be the case that there is nectar in both locations. There is no capacity for one bee dance to *dispute* another one. In her *Language, Thought and Other Biological Categories*, Millikan (1984) argues that explicit contradiction relies on subject-predicate structure and is thus attendant to our capability to communicate propositional content, i.e., that without it conceptual symbolic language is not possible.

Future Work

Independent of the issues raised above there remains scope to explore the extent to which simple CTRNN-controlled agents can be evolved to solve increasingly complicated signalling tasks. Future work of this kind could, for example increase the *number of agents* involved in each signalling interaction (beyond the pair of agents considered in the literature so far), or increase the *number of tasks* for which the agents are required to employ communication (beyond

conveying the location of a target). Honey bee dances, for instance are performed in front of a large number of observers, each of whom integrates the information from several dances before flying towards one (presumably the best) of the advertised locations (von Frisch, 1965). Likewise, honey bees are able to use their dance language to communicate about more than the location of nectar-bearing flowers since they also use the signalling system to decide amongst competing nest sites (Lindauer, 1955; Seeley & Visscher, 2004). Extensions of this kind would further separate signalling from acting, creating an intervening gap that may require more complex cognitive mechanisms than have been displayed in the literature so far.

Whether advertising nectar or nest sites, the honey bee dance language is being used to solve an inherently *analogue* problem: communicating the relative distance and direction to target locations in continuous space. This type of task seems inherently suited to a CTRNN architecture that is also inherently analogue and also to the setting of simple simulated robots moving in a one- or two-dimensional arena. One issue is that evolving and analysing CTRNN agent solutions to such tasks may reveal more about the behavioural biases and capacities of the CTRNN substrate, rather than the nature of the cognition or communication required by the task (Bullock & Cliff, 1997). A second, more significant problem is that such analogue tasks can also encourage a kind of ‘faux displacement’. Whilst an agent that is limited to communicate in one region of space before navigating alone to a target location that is temporally and physically distant may be described as exploiting information in a signal that is ‘about’ this displaced location, it can be equally accurate to describe this communication without invoking displacement at all. Such an agent can be re-described as having received a signal that is ‘about’ the way that it should start moving *right here* and *right now*, behaviour that merely has the *side-effect* of ensuring arrival at the right location later on. As such, there may be value in exploring signalling scenarios that lie outside the set of analogue spatial navigation problems, scenarios in which communication about *discrete* and *articulated* referents is required (e.g. communicating the identity of target objects that differ in multiple qualitative ways, Steels, 2003) particularly if the referents need to be acted on in the future in some way that is determined by the context in which they are encountered.

For instance a task in which an agent must combine multiple sources of information about the identity of a target, and must interact with that target when they encounter it in the future in a way that is determined by contextual factors would represent a significant departure from tasks explored in the literature to date (e.g. being told “The traitor wears a hat” + “...has blonde hair” + “...has a big nose” might allow me to identify them during a sequence of interactions with suspects, but whether to expose

the traitor or keep quiet might depend on whether I am in love with them or not).

Evolving agents to successfully solve more demanding signalling challenges such as these will require more sophistication on the part of the evolved agents and should encourage strategies that are not as reliant on the close coupling between signaller and receiver that was central to the solutions reported here. It is not a coincidence that, simultaneously, studies of this type would address directly several of the items on Millikan’s list of representational properties presented above.

Conclusion

A single, three-node continuous time recurrent neural network was evolved successfully to solve a task requiring two agents controlled by the same network to communicate private information to each other and to follow a course of action that depended on a comparison between these two pieces of information. The evolved solution relied on a signalling channel constructed from the physical movement of the agents as detected by their simple proximity sensors.

Whilst the evolved behaviour is an example of bidirectional referential communication of a displaced referent, it is not easily described in terms of information processing (i.e. the storage and processing of structure-sensitive symbolic representations), being instead more naturally explained in dynamical systems theoretic terms as relying on a carefully parameterised period of structural coupling between the two agents. The evolved signalling system is reminiscent of that employed by honey bees in that it exhibits only limited productivity and referent displacement, and employs signals that are analogue and non-compositional, with self-referring elements and an inseparably indicative-imperative mood. As such, it lacks several representational properties that are the hallmarks of human language. In order to achieve more of these language-like features in an evolutionary simulation model, it appears likely that the research paradigm being employed here may need to move beyond communication tasks that require agents to direct one another to locations within a continuous co-ordinate space.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by funding from the UK Engineering and Physical Sciences Research Council Grant Award EP/R004757/1 entitled *T-B PHASE Prosperity Partnership*.

ORCID iD

Seth Bullock  <https://orcid.org/0000-0003-4371-5965>

References

- Artiga, M. (2014). Signaling without cooperation. *Biology and Philosophy*, 29(3), 357-378. <https://doi.org/10.1007/s10539-014-9436-0>
- Baker, J. E. (1987). Reducing bias and inefficiency in the selection algorithm. In J. J. Grefenstette (Ed), *Proceedings of the second international conference on genetic algorithms and their application (ICGA'87)* (pp. 14-21). Cambridge, MA: Lawrence Erlbaum Associates.
- Beer, R. D. (1995). On the dynamics of small continuous-time recurrent neural networks. *Adaptive Behavior*, 3(4), 469-509. <https://doi.org/10.1177/105971239500300405>
- Beer, R. D. (2000). Dynamical approaches to cognitive science. *Trends in Cognitive Sciences*, 4(3), 91-99. [https://doi.org/10.1016/s1364-6613\(99\)01440-0](https://doi.org/10.1016/s1364-6613(99)01440-0)
- Bullock, S. (1997). *Evolutionary simulation models: On their character, and application to problems concerning the evolution of natural signalling systems*. PhD Thesis. UK: University of Sussex.
- Bullock, S. (1998). A continuous evolutionary simulation model of the attainability of honest signalling equilibria. In C. Adami, R. Belew, H. Kitano, & C. Taylor (Eds), *Artificial life VI: Proceedings of the sixth international conference on the synthesis and simulation of living systems* (pp. 339-348). Cambridge, MA: MIT Press.
- Bullock, S. (2004). Making room for representation. *Adaptive Behavior* 11(4): 279-280. Commentary on RD Beer's "The Dynamics of Active Categorical Perception in an Evolved Model Agent", *Adaptive Behavior*, 11(44), 209-243.
- Bullock, S. (2008). Charles Babbage and the emergence of automated reason. In P. Husbands, O. Holland, & M. Wheeler (Eds), *The mechanical mind in history* (pp. 19-39). Cambridge, MA: MIT Press.
- Bullock, S. (2009). In defence of the abstracted animat. *Adaptive Behavior* 17(4): 303-305. Commentary on B Webb's "animals versus animats: Or why not model the real iguana?". *Adaptive Behavior*, 17(4), 269-286.
- Bullock, S. & Cliff, D. (1997). The role of 'hidden preferences' in the artificial co-evolution of symmetrical signals. *Proceedings of the Royal Society of London, Series B: Biological Sciences*, 264(1381), 505-511. <https://doi.org/10.1098/rspb.1997.0072>
- Campos, J. I., & Froese, T. (2017). Referential communication as a collective property of a brain-body-environment-body-brain system: A minimal cognitive model. *IEEE symposium series on computational intelligence (SSCI 2017)* (pp. 1-8). IEEE.
- Campos, J. I., & Froese, T. (2019). From embodied interaction to compositional referential communication: A minimal agent-based model without dedicated communication channels. In H. Fellersmann, J. Bacardit, A. Goñni-Moreno, & R. M. Fuchslin (Eds), *ALIFE 2019: The 2019 conference on artificial life* (pp. 79-86). MIT Press.
- Fodor, J. A. (1975). *The language of thought*. Harvester Press.
- Fodor, J. A., & Pylyshyn, Z. W. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition*, 28(1-2), 3-71. [https://doi.org/10.1016/0010-0277\(88\)90031-5](https://doi.org/10.1016/0010-0277(88)90031-5)
- Johnstone, R. A. (1997). The evolution of animal signals. In J. R. Krebs & N. B. Davies (Eds), *Behavioural ecology: An evolutionary approach* (4th edition, pp. 155-178). Oxford: Blackwell.
- Kirby, S. (2002). Natural language from artificial life. *Artificial life*, 8(2), 185-215. <https://doi.org/10.1162/106454602320184248>
- Krebs, J. R., & Dawkins, R. (1984). Animal signals: Mind reading and manipulation. In J. R. Krebs & N. B. Davies (Eds), *Behavioural ecology: An evolutionary approach* (2nd Edition, pp. 380-402). Oxford: Blackwell.
- Lindauer, M. (1955). Schwarmbienen auf Wohnungssuche. *Zeitschrift für vergleichende Physiologie*, 37(4), 263-324. <https://doi.org/10.1007/bf00303153>
- MacLennan, B., & Burghardt, G. (1993). Synthetic ethology and the evolution of cooperative communication. *Adaptive Behavior*, 2(2), 161-187. <https://doi.org/10.1177/105971239300200203>
- Manicka, S. (2012). Analysis of evolved agents performing referential communication. In C. Adami, D. M. Bryson, C. Ofria, & R. T. Pennock (Eds), *ALIFE 2012: The thirteenth international conference on the synthesis and simulation of living systems* (pp. 393-400). Cambridge, MA: MIT Press.
- Matarić, M., & Cliff, D. (1996). Challenges in evolving controllers for physical robots. *Robotics and Autonomous Systems*, 19(1), 67-83. [https://doi.org/10.1016/s0921-8890\(96\)00034-6](https://doi.org/10.1016/s0921-8890(96)00034-6)
- Maynard Smith, J., & Harper, D. (2003). *Animals signals*. Oxford University Press.
- Millikan, R. G. (1984). *Language, thought and other biological categories*. Cambridge, MA: MIT Press.
- Millikan, R. G. (1989a). Biosemantics. *The Journal of Philosophy*, 86(6), 281-297. <https://doi.org/10.2307/2027123>
- Millikan, R. G. (1989b). In defense of proper functions. *Philosophy of Science*, 56(2), 288-302. <https://doi.org/10.1086/289488>
- Noble, J., Di Paolo, E. A., & Bullock, S. (2001). Adaptive factors in the evolution of signalling systems. In A., Cangelosi & D. Parisi (Eds), *Simulating the evolution of language* (pp. 53-78). Heidelberg: Springer.
- Parker, G. A. (1974). Assessment strategy and the evolution of animal conflicts. *Journal of Theoretical Biology*, 47(1), 223-243. [https://doi.org/10.1016/0022-5193\(74\)90111-8](https://doi.org/10.1016/0022-5193(74)90111-8)
- Quinn, M., (2001). Evolving communication without dedicated communication channels. In J. Kelemen, & P. Sosik (Eds), *Advances in artificial life: Sixth European conference on artificial life (ECAL'01)* (pp. 357-366). Berlin: Springer.

- Seeley, T. D., & Visscher, P. K. (2004). Group decision making in nest-site selection by honey bees. *Apidologie*, 35(2), 101-116. <https://doi.org/10.1051/apido:2004004>
- Steels, L. (1997). Self-organising vocabularies. In C. G. Langton & K. Shimohara (Eds), *Artificial life V: Proceedings of the fifth international workshop on the synthesis and simulation of living systems* (pp. 179-184). Cambridge, MA: MIT Press.
- Steels, L., (2003). Evolving grounded communication for robots. *Trends in Cognitive Sciences*, 7(7), 308-312. [https://doi.org/10.1016/s1364-6613\(03\)00129-3](https://doi.org/10.1016/s1364-6613(03)00129-3)
- Van Gelder, T., & Port, R. F. (1995). It's about time: An overview of the dynamical approach to cognition *Mind as motion: Explorations in the dynamics of cognition* (pp. 1-43). MIT Press.
- von Frisch, K. (1965). *The dance language and orientation of bees*. Harvard University Press.
- Werner, G. M., & Dyer, M. G. (1992). Evolution of communication in artificial organisms. In C. G. Langton, C. Taylor, J. D. Farmer, & S. Rasmussen (Eds), *Artificial life II* (pp. 659-687). Redwood City, CA: Addison-Wesley.
- Williams, P. L., Beer, R. D., & Gasser, M. (2008). Evolving referential communication in embodied dynamical agents. In S. Bullock, J. Noble, R. Watson, & M. Bedau (Eds), *Artificial life XI: Proceedings of the eleventh international conference on the simulation and synthesis of living systems* (pp. 702-709). MIT Press.

About the Authors



Ruairi Fox is a recent graduate from the department of Computer Science at the University of Bristol. He received a Masters of Engineering in Computer Science from the University of Bristol in 2021. His research interests include evolutionary robotics and minimal cognition as well as artificial intelligence as a whole.



Seth Bullock was appointed as Toshiba Chair in Data Science and Simulation at the University of Bristol in 2015. His expertise is in complex systems simulation modelling for adaptive systems. He has led complex systems simulation modelling projects on social care, infrastructure robustness and resilience, and design principles for human-robot teams. He served as Conference Chair for ALIFE XI (2008), was twice elected to the Board of the International Society for Artificial Life (2008–2012, 2012–2016), has consulted for the UK Government on complexity in ICT (2004) and financial systems (2012), is a fellow of the European Centre for Living Technology (2022-), and publishes in journals spanning computing, engineering, and the social and life sciences.