DOI: http://dx.doi.org/10.7551/978-0-262-31709-2-ch074

Hidden information transfer in an autonomous swinging robot

James Thorniley[1] and Phil Husbands[1]

[1]Centre for Computational Neuroscience and Robotics, Dept. of Informatics, University of Sussex, UK, BN1 9QG
jt241@sussex.ac.uk

Abstract

This paper describes a hitherto overlooked aspect of the information dynamics of embodied agents, which can be thought of as hidden information transfer. This phenomenon is demonstrated in a minimal model of an autonomous agent. While it is well known that information transfer is generally low between closely synchronised systems, here we show how it is possible that such close synchronisation may serve to "carry" signals between physically separated endpoints. This creates seemingly paradoxical situations where transmitted information is not visible at some intermediate point in a network, yet can be seen later after further processing. We discuss how this relates to existing theories relating information transfer to agent behaviour, and the possible explanation by analogy to communication systems.

Introduction

The dynamics of embodied agent-environment systems are increasingly analysed using information theory (Lungarella and Sporns, 2006; Pfeifer et al., 2007b; Bertschinger et al., 2008; Klyubin et al., 2008; Pitti et al., 2009; Williams and Beer, 2010; Moioli et al., 2012; Schmidt et al., 2012). This paper adopts this approach and demonstrates a phenomenon that is consistent with the analogy to communications, but thus far seemingly overlooked in studies of information transfer in embodied agents. We describe "hidden" information transfer in a simulated robot: strongly physically coupled parts of the system carry information between separated endpoints, without such information tranfer being visible between the carrier components themselves.

Information transfer is often characterised using forms of *transfer entropy* (Schreiber, 2000), itself a nonlinear generalisation of Granger causality (Barnett, 2009). Information transfer from X to Y is quanitified by the relative improvement in statistical prediction of the future states of Y when the current state of X is known in addition to the already-known historical states of the target variable Y. A known issue, and potential source of confusion, is that information transfer is not a measure of the *physical strength* of coupling – a common example being synchronised systems, where very high coupling may mean that two time series are almost identical, leading to little prediction improvement and hence low transfer entropy in spite of strong physical coupling (this is seen in e.g. Thorniley, 2011). This is sometimes regarded as a failure of transfer entropy to properly capture *causal* influences (Ay and Polani, 2008; Lizier and Prokopenko, 2010; Janzing and Balduzzi, 2012). The agent model used in this paper will exhibit this type of phenomenon, but in addition we will show that although strongly coupled components may exhibit low transfer entropy, they may still act as information conduits, hiding information transfer between more separate components.

Our model is a reactive robot designed to behave like a child swinging on a swing. As the feedback gain in the robot's controller increases, a self sustaining oscillation is created. The agent has a simple neural model acting as its brain, which is connected to the environment via its body. The state of the agent's neural system cannot (physically) influence the environment apart from by first affecting its body. However, we demonstrate that information transfer can take place from brain to environment *without* information transfer from brain to body. This shows how information transfer can be hidden within the agent, and revealed by its interaction with the environment.

This is the key result of this paper – information can pass through a chain of coupled systems, e.g. A to B to C such that there is a high information transfer from A to C but *not* from A to B, even though physically there is no alternative route. In the discussion at the end of the paper we will consider how similar effects occur in communication systems by way of analogy to our agent based model.

This paper is organised as follows: the next section below describes the model swinging agent and its general dynamical features. The analysis in the following section shows the information hiding phenomenon by analysing the information transfer between each component of the system. The final section discusses this result and considers the implications for the study of embodied autonomous agents in terms of information theory.

Table 1: Variables and parameters

Symbol	Type/Value	Description
θ	Variable	Angle of pendulum from downward vertical
ω	Variable	Angular velocity of pendulum ($d\theta/dt$)
r	Variable	Current pendulum extension
v	Variable	Rate of pendulum extension (dr/dt)
u	Variable	Force control variable – force on bob due to effector
F_a	Intermediate	Force on bob due to acceleration
F_s	Intermediate	Force on bob due to spring
A	Independent variable (0-80)	Motor neuron output at saturation
g	9.81	Acceleration due to gravity
b	0.3	Pendulum damping coefficient
ρ	2	Motor neuron sensitivity
ϕ	20	Control parameter
k	100	Spring force constant
c	20	Spring damping ($= 2\sqrt{k}$ for critical damping)

Reactive swinging agent

The system studied here is a simplified model of a child swinging on a swing. The swing itself will be modelled as a rigid massless rod attached to a fixed pivot at one end with a mass (being the mass of the agent) at the other end. The agent's motor control consists in its ability to move the mass up and down (towards and away from the pivot). There are two general ways to approach the dynamical modelling of such a system. It is possible to use a "kicked pendulum" approach where a periodic forcing function is used to perturb the mass (e.g. Belyakov et al., 2009). However it has been found that even though a pendulum can be made to swing this way, the limit cycle produced is in fact unstable, and thus this is in a practical sense impossible to achieve in the real world, suggesting that a better approach is the "self-excited" oscillator (Pinsky and Zevin, 1999; Zevin and Filonenko, 2007). Here the agent creates a positive feedback loop by adjusting the distance from the mass to the pivot point (e.g. by raising and lowering the centre of mass of the agent relative to a fixed attachment point at the end of the rod). This create a stable limit cycle as well as a resting point (where the swing is pointing straight downwards and there are no vibrations to amplify). Thus if the swing is given an initial "push", the movement of the agent will sustain the oscillation, hence the system is described as self-excited. This approach treats the agent as a reactive system in the sense of Brooks (1986). This section provides further details on the implementation of this system.

A representation of the model is shown in figure 1. There is a massless rod with length normalised to one arbitrary unit. It makes an angle θ with the vertical axis along which the gravitational force g applies. The "agent" consists of a mass-spring-damper system attached to the end of the rod. The mass is influenced by the gravitational force, along with

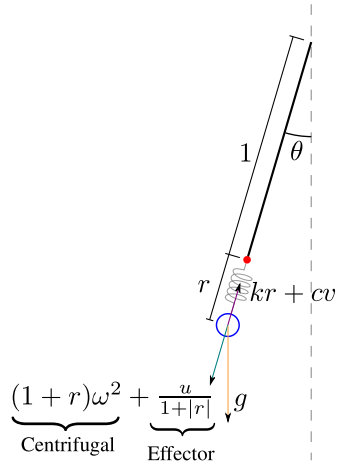


Figure 1: Spring based model of the swinging agent

the centrifugal effect of rotation and the forces created by the spring: linear contraction kr where k is a constant and r is the extension of the spring, and damping cv with c another constant and $v = \dot{r}$ – the linear velocity of the mass in the direction of the spring. The agent creates an effector force u which acts on the mass, but this is derated according to the current absolute extension of the spring, modelling a linear motor which produces less force output when it is already extended.

The full system can be described by the following equations. Table 1 lists each of the variables and parameters used. Dots represent differentiation with respect to a non-dimensionalised time variable t[1]:

[1]For simplicity all variables are treated as dimensionless, though the choice of $g = 9.81$ suggests the system could be treated as a one metre long pendulum with the agent mass at one kilogram, and time in seconds.

$$\dot{\theta} = \omega \quad (1)$$
$$\dot{\omega} = -\frac{g}{1+r}\sin(\theta) - b\omega \quad (2)$$
$$\dot{r} = v \quad (3)$$
$$\dot{v} = \frac{u}{1+|r|} + F_a + F_s \quad (4)$$
$$\dot{u} = \phi(A\tanh(\rho v) - u) \quad (5)$$

The last equation describes the internal dynamics of the agent's reactive controller. The agent senses the current velocity v of its spring, and passes this through a simple sigmoidal neuron, which determines a desired output force $A\tanh(\rho v)$ where A and ρ are parameters. The actual output force u moves towards this desired value in proportion to its current error according to the rate parameter ϕ.

The acceleration of the mass in the direction of the pendulum rod $\dot{v}$ is given by the resultant force (we assume the mass is normalised to one arbitrary unit). That is, equation 4 shows $\dot{v}$ is the sum of the force due to acceleration F_a (i.e. gravity and centrifugal forces, equation 6) and the force due to the spring F_s (equation 7) along with the effector force described above.

$$F_a = g\cos(\theta) + (1+r)\omega^2 \quad (6)$$
$$F_s = -kr - cv \quad (7)$$

We can treat the different dynamical variables as components of either the agent or environment, and further subdivide the agent into "brain" and "body" as shown in figure 2. The intention is to treat the agent as dynamical system which is "embodied" in the sense that its overall behaviour is a result of the close coupling of the agent's body, brain and environment (Pfeifer et al., 2007a). The main sensor variable is v – the input to the neuron, though the spring extension r can also be conceptualised as a component of the agent's sensory system. The motor output is represented by u, and the environment consists of the pendulum system: ω and θ.

The fixed parameter values used in the following simulations are shown in table 1. The parameter A effectively controls the feedback gain and will be varied as the independent variable in what follows.

The bifurcation plot in figure 3 gives an indication of the general dynamical features of the system. These plots are obtained by recording the angular speeds at which the swing passes through the downward direction, having been intialised with a random angular velocity and the "transient" time while the system is still far from a stable cycle or point discarded. Data is obtained using Runge-Kutta integration – all results in this paper are based on an integration step size of 1/50th of a time unit, with a simulation length of 1000 time units. With A low, less than about 10, there is a

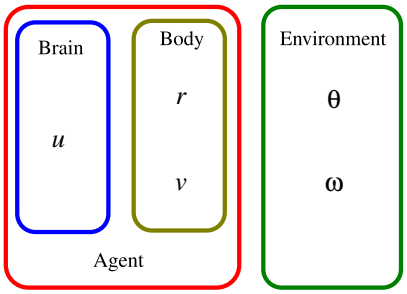


Figure 2: The agent and environment in terms of dynamical variables

single, globally stable fixed point – i.e. there is insufficient feedback for the agent to actually swing. Between feedback gains of around 10 and around 50, the agent usually swings side to side (represented in blue in the figure) – where the agent returns to $\theta = 0$ swinging in a different direction each time. Above $A = 30$ another stable cycle appears where the pendulum swings over the top rather than side-to-side, i.e. it returns to $\theta = 0$ travelling in the same direction (same sign of ω) each time. Note that the two cycles coexist between values of A around 30 to 50, but above that only the rotating motion occurs. Finally, above $A = 70$, a transition to chaotic motion occurs – above this point the system will sometimes rotate and sometimes swing side to side during a single trajectory. The fixed point where the system does not swing is locally stable for values of A less than around 33, meaning that sometimes the system will tend towards resting rather than either of the limit cycles. Thus the ultimate behaviour of the system is in general dependent on the initial conditions as well as the particular value of A chosen.

We now consider a slight alteration to the model. In practice, no sensor is perfect, and thus the input to the neuron might conceivably be modelled as a stochastic variable with a slight perturbation ε_v, so equation 5 becomes:

$$\dot{u} = \phi(A\tanh(\rho(v + \varepsilon_v)) - u)$$

Assuming ε_v is small we can linearise its effect model it as a random additive perturbation on $\dot{u}$:

$$\dot{u} = \phi(A\tanh(\rho v) - u) + \phi A\rho\,\mathrm{sech}^2(\rho v)\varepsilon_v$$

In order to practically simulate the system, it must be written as stochastic differential equations. Specifically, we convert the equation for u (which is the only variable where we directly add noise) into *Langevin equation* form:

$$du = \phi(A\tanh(\rho v) - u)dt + \phi A\rho\,\mathrm{sech}^2(\rho v)\sigma dW$$

Where W represents a Wiener process, and a new parameter σ is introduced to control the strength of the random

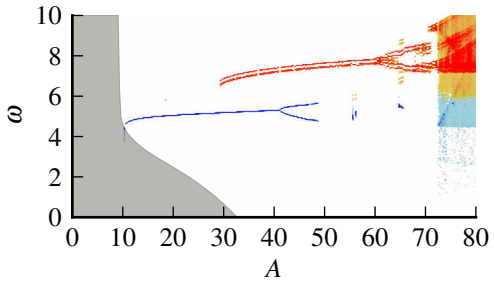


Figure 3: Bifurcation plot showing the behaviour of the system as the internal gain A is increased. Simulations are performed at each of 300 linearly spaced points between $A = 0$ and $A = 80$. Plot shows absolute angular velocity of the pendulum recorded as it passes through the "downward" ($\theta = 0$) plane of its state space – points in blue are returns to $\theta = 0$ where the sign of ω changed in between returns (i.e. the agent is swinging side to side) and points in red show returns in the same direction (the pendulum has swung over the top). The grey area shows the numerically estimated stability region for the fixed point where $\omega = 0$ – i.e. if the system is within this region it will eventually stop swinging. Outside of this region, it will go to either a swinging or rotating stable cycle.

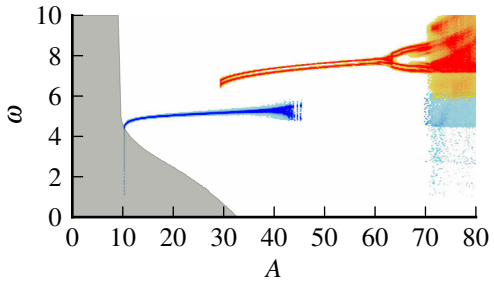


Figure 4: Bifurcations with noise $\sigma = 0.25$

noise. This equation can be numerically solved using the stochastic strong order 1.0 Runge-Kutta algorithm. The full details of this approach including integration algorithm are found in Sauer (2012). The overall effect is that u behaves as if the neuron senses the current velocity v with additive Gaussian white noise, where the noise power is increased by increasing the newly introduced parameter σ. Figures 4 and 5 show the effect of increasing σ on the bifurcation structure – the main features remain much the same, but the crossing points are now somewhat random.

As well as making the model more "realistic", introducing this random perturbation ensures that the system is generally ergodic, which facilitates the correct calculation of transfer entropy. Without this property, the probabilities estimated from time series data tend to make little sense (see Breiman,

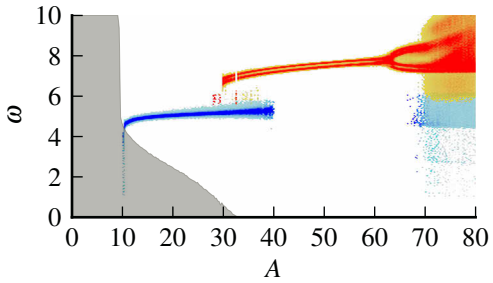


Figure 5: Bifurcations with noise $\sigma = 0.5$. Note that since the system is stochastic, the stability regions are not deterministically defined (close to the edge of the region shown, some trajectories may tend towards the fixed point and some towards the limit cycle depending on chance). The region shown shaded corresponds to the median stable boundary found in 20 simulation runs at each value of A.

1969, for a discussion). This slight randomness also means that even for very closely synchronised variables there will likely be at least some transfer entropy measured, as there will be a constant introduction of entropy inside the system.

Information transfer analysis

Transfer entropy is generally defined for two time series X and Y as a relative entropy or conditional mutual information:

$$TE_{X\to Y} = \sum P(x_t, y_{t+\delta}, y_t) \log \frac{P(y_{t+\delta}|x_t, y_t)}{P(y_{t+\delta}|y_t)}$$

The data points being taken at discrete time intervals δ, e.g. $X = (x_{t0}, x_{t0+\delta} \ldots x_{t0+n\delta})$. The sum is taken over the support of $P(x_t, y_{t+\delta}, y_t)$ – i.e. all possible combinations of values for the three variables. In this analysis we use the time interval $\delta = 1$ (i.e. 50 integration steps, corresponding to approximately one quarter of a cycle).

It is problematic to calculate the transfer entropy on continuous-valued time series such as we have here. We have used symbolic transfer entropy (Staniek and Lehnertz, 2008), which uses a convenient rank transform to find an estimate of the transfer entropy on continuous data without the need for kernel density estimation.[2] First an embedding dimension m is chosen (we use 4), for each $n \geq m$ we set $\hat{x}_{t0+n\delta} = \mathrm{rank}[(x_{t0+(n-m+1)\delta} \ldots x_{t0+n\delta})]$, where rank converts a sequence into its sort order, e.g. $(0.0, 0.4, 0.3, 0.25)$ becomes $(1, 4, 3, 2)$. That is, each original observation (after embedding in m dimensions) is a continuous vector ($x_t \in \Re^m$) and after transformation each observation is assigned one of the $m!$ possible permutations

[2] Alternatives exist such as k nearest-neighbour methods (Kraskov et al., 2004; Evans, 2008). At this time we are not aware of a reason to prefer one method over the other in this instance.

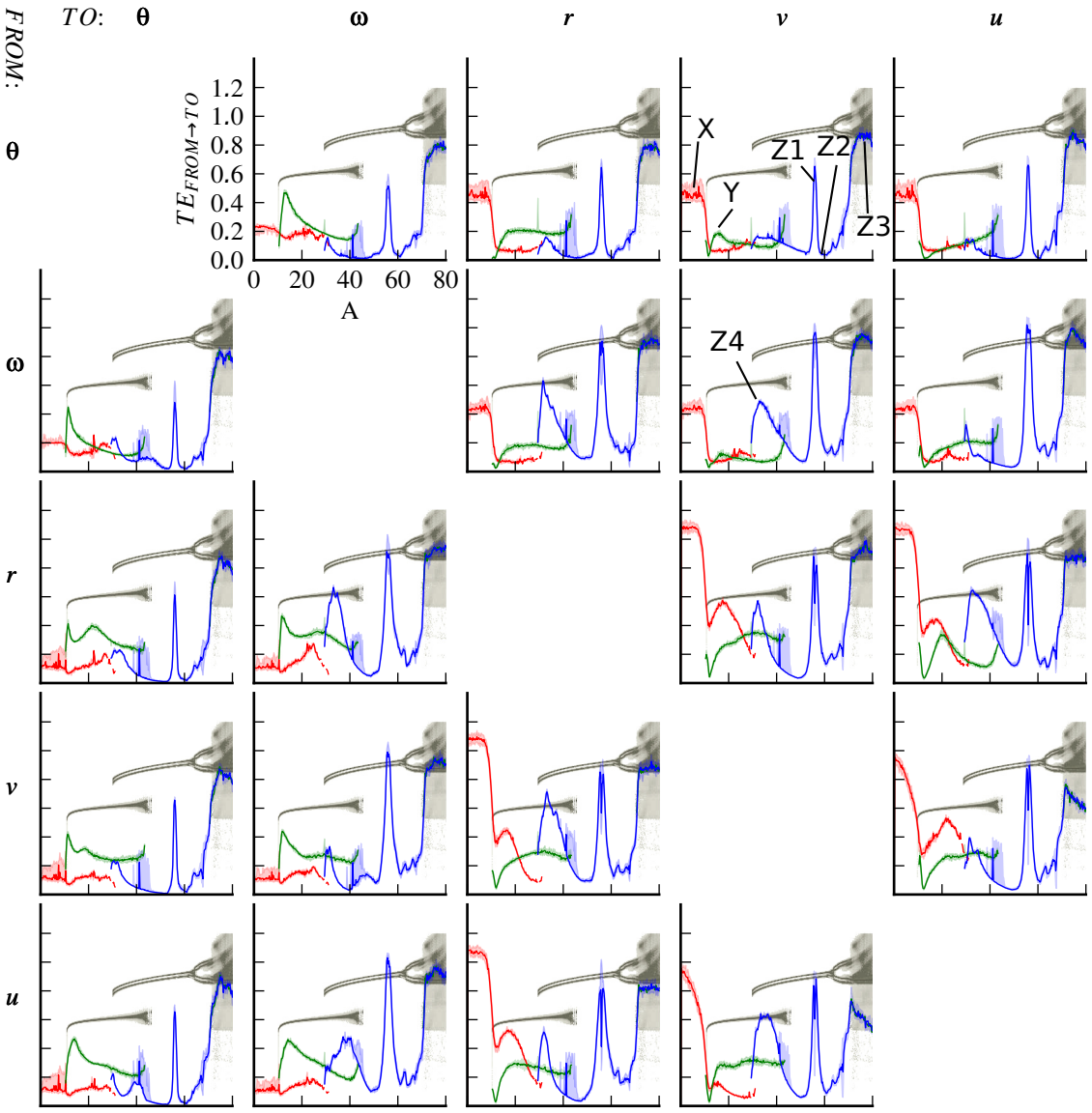
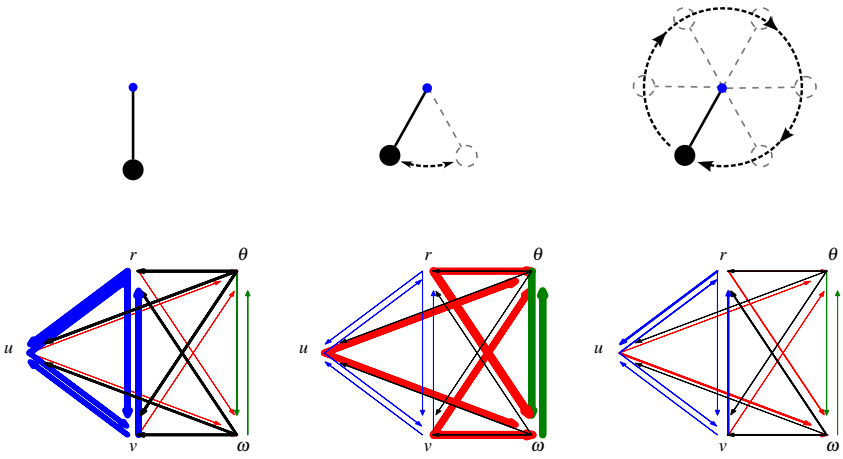


Figure 6: Symbolic transfer entropy in bits from each dynamical variable to each other one as the internal gain A is varied in the system with noise $\sigma = 0.25$. Note that figure 7 shows some of the same data in a form that is easier to interpret for the effects we are primarily interested in – the current figure is provided to show the context for the particular values of A chosen for re-plotting in figure 7. The background of each plot shows in grey a copy of the bifurcation diagram from Figure 4 – this is intended to help identify the correspondence between recorded transfer entropy and system behaviour. The results from 20 runs are shown after grouping by behaviour mode (color online): red for stable (non-swinging), green for side-to-side swinging and blue for rotational motion. For each behaviour the median is calculated for plotting and the shaded area around each line shows the 10th-90th percentile range where it is visible (for most values of A there was very little variation in the results). For comparison, the bifurcation plot for the system is shown in grey in the background. Some key points on the graphs are labelled in the $\theta \to v$ and $\omega \to v$ plots (the same features are present on some of the other plots as can be seen): at X the transfer entropy for low feedback gains (stable behaviour) is often high; at Y there is a peak in the curve for side-to-side swinging behaviour at around $A = 12$; Z1 is a notable peak in the rotational swinging behaviour, which appears to correspond to some complexity in the behaviour not captured by the bifurcation diagram, Z2 is a trough at around $A = 50$; Z3 and Z4 show peaks in transfer entropy which can easily be related to features of the bifurcation diagram which indicate higher complexity – the chaotic behaviour at very high gains and the behaviour close to the bifurcation point.



(a) Stable (b) Swinging (c) Rotating

Figure 7: Median transfer entropy under three different behavioural regimes represented by arrow widths. Arrows are colored blue for information transfer within the agent (variables u, r and v), green for within the environment (variables θ and ω), red for agent to environment and black for environment to agent. See Figure 2 for classification of variables. (a) Low feedback gain ($A = 2.6$), the system cannot maintain a periodic motion and tends towards the stable state. Higher transfer entropy is seen within the agent. (b) Moderate gain ($A = 12$), the agent will swing side to side. This graph illustrates the information hiding effect (see text). (c) High feedback ($A = 50$), the system rotates over the top. At this value almost no transfer entropy is seen in any direction. Note that the arrow widths in (a) are 1/3rd the scale of the widths in (b) and (c) since the transfer entropy values are generally much larger in (a).

of a sequence of lenth m. The permutation is denoted $\hat{x}_t$ and for ease of calculation could obviously be assigned an integer representation according to an arbitrary one-to-one mapping. The formula for symbolic transfer entropy is then

$$STE_{X\to Y} = \sum P(\hat{x}_t, \hat{y}_{t+\delta}, \hat{y}_t) \log \frac{P(\hat{y}_{t+\delta}|\hat{x}_t, \hat{y}_t)}{P(\hat{y}_{t+\delta}|\hat{y}_t)}$$

With the probabilities estimated in the natural manner for discrete variables according to frequency of occurrence, i.e. $P(\hat{x}_t = X)$ would simply be the number of time points where $\hat{x}_t$ is found to be X divided by the total number of observations taken.

On every experimental run, the system is initiated with all dynamical variables set to zero except for ω which is taken uniformly at random from $[-10, 10)$. The first 100 time units are treated as transient non-stationary data and discarded, and the remaining 900 data points are fed to the symbolic transfer entropy calculation. This process is repeated ten times with different initial conditions, and the trajectories recorded are classified according to their final behaviour mode: resting, swinging or rotating.

The set of results in figure 6 shows all the transfer entropy values calculated for the system using a noise amplitude of $\sigma = 0.25$, taking each possible combination of source and target variables. This shows a few basic features of the results. We see as expected that the transfer entropy does not straightforwardly correspond to physical coupling – there is no simple correspondence between the independent variable A and the transfer entropy value. We also see that very different patterns of information transfer are observed for the different behavioural regimes, even at the same value of A.

A simpler graphical representation of the transfer entropy is shown in figure 7. This shows the median transfer entropy for a particular behaviour at a chosen value of A as the width of an arrow pointing in the direction of information transfer. The arrows have been colour coded by the way in which they connect the brain, body and environment components.

The most striking result for our purposes is shown in figure 7b, where the feedback gain is moderate, resulting in a natural swinging behaviour. Here, the highest information transfer is along the paths coloured red which emanate from the agent (according to the classification in Figure 2) and flow towards the environment. This includes the arrows which directly connect the output of the motor neuron u to the environment variables θ and ω. However, there is no direct physical connection along this path since the coupling between the brain and environment is always mediated by the body. This is shown in equations 1 to 5 – the neuron output u does not appear on the right hand side of the equations for $\dot{\theta}$ and $\dot{\omega}$, and hence it can only influence these variables through the intermediate coupling to its body (since the body displacement r *does* influence ω). Thus the informa-

tion transferred from u to ω (shown by a thick red arrow) for example is surely carried across the chain $u \to v \to r \to \omega$, yet there is *low* information transfer from u to v and r (illustrated by the thin blue arrows). It is in this sense that we claim this shows a form of hidden information transfer – we know that the brain can only influence the environment by going through the body, but even when a high information transfer is measured from brain to environment, there is a smaller amount from brain to body

Figures 7a and 7c do not clearly show this phenomenon, since it is in no sense necessary for it to be present. Figure 7a seems to show the strongest connections within the agent when the feedback gain is low and the system is resting, which can be explained by the fact that the source of entropy here is the sensor noise inside the agent, and since the agent is not swinging it may move up and down, but is not likely to influence the angle of the pendulum. In figure 7 the very high feedback coupling is likely creating a highly synchronised dynamic where the observed transfer entropy is very low.

Discussion

The key result of this work is shown in figure 7b, where during the entrained oscillatory motion of the system, the transfer entropy is shown to be higher from the brain to the environment than it is from the brain to the body, even though it is not possible for the brain to influence the environment without that influence passing through the body.

It appears that the entrained behaviour leads to a reduction in the transfer entropy measured within the agent, as can be seen by comparing the blue arrows between figures 7a and 7b. This is likely due to the close synchrony between these variables when the agent is swinging – a factor that is known to generally reduce measured transfer entropy. What is interesting is that though the swinging behaviour appears to decrease the transfer entropy within the agent, it also corresponds to increased information transfer from the agent to its environment. This is a clear demonstration of the importance of the agent's embodiment to the information dynamics of the system – the interesting (as in measurable) interaction takes place between the agent and the environment rather than within the agent.

What we are calling *information hiding* is the way in which information coming from a variable we specifically associate with the agent's neural system, i.e. u, appears to pass straight to the environment without having to "go through" the body, in spite of that the fact that we already know that, in a physical sense, it must, since only the agent's body is physically coupled to the environment.

It is worth attempting to gain a little intuition for how this effect is working. For an analogy that is perhaps useful in the current context, consider the simplest type of encryption system based on a symmetric key illustrated in figure 8. A key is a randomly chosen binary sequence that

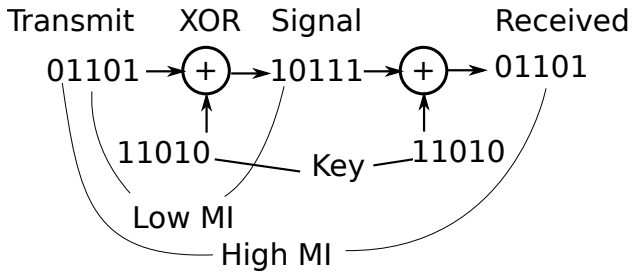


Figure 8: A simple encryption system

has been previously shared between a sending and receiving party. The sender can encrypt a message by performing the XOR operation bit-wise between the key and the message. However, since the key was chosen randomly, the resulting encrypted signal should be statistically independent of the transmitted message – the encryption operation appears (to anyone without the key) to flip bits of the message at random (i.e. it randomly changes some 1's to 0's and some 0's to 1's). However, with the key, it is trivial to reconstruct the original message – the same XOR operation is simply applied using the previously shared key. Symbolically, if we have a transmitted message T, encrypted signal S and received signal R then we have a very low $I(T;S)$ yet high $I(T;R)$. Though expressed in terms of mutual information rather than transfer entropy, this is essentially the same information hiding phenomenon as we have been discussing. Indeed if we assume the individual bits of the message and the key are independent of each other then $TE_{T\to S} = I(T;S|\text{history}(S)) = I(T;S)$ and so on.

The information hiding process can thus be seen as a message being obscured by at some point and later reconstructed. In the example above this function is performed by the encryption system and is dependent on having a piece of secondary data (the key) shared between the two endpoints via some alternative channel to the main signal path. Of course, the encryption system is carefully designed to achieve this – it requires the deliberate sharing of the key. However, comparable processes have been found relying only on chaotic synchronisation: Cuomo and Oppenheim (1993) demonstrated that synchrony between a pair of coupled Lorenz attractor systems can be used to "hide" information in a similar way.[3] Their experiment suggests that it is plausible that information could be hidden by a dynamical process such as the one studied here without the need for the deliberate design of an encryption system.

This phenomenon should not be viewed as information being completely lost to the world and then coming back

[3]Note that this system is not generally regarded as computationally secure as an encryption mechanism since the reconstruction circuit (which effectively serves as the "key") can be relatively easily inferred using attractor reconstruction on the transmitted signal.

– rather it is simply hidden and then reconstructed by the action of some dynamical system. We have interpreted information here as a statistical summary of collected data – not as a physical quantity that exists in the world.

We have said little explicitly about causation, though of course to say that the brain must influence the environment via the body suggests a causal interpretation. Recent work has studied the relationship between transfer entropy and causal inference in part motivated by phenomena similar to the one described here (e.g. Ay and Polani, 2008; Lizier and Prokopenko, 2010). Information theory has also been applied successfully in the context of embodied systems (e.g. Ay et al., 2008; Klyubin et al., 2008). Both of these connections are relevant: can information hiding as presented here be useful in any sense as a guide to causal inference? How should the current case study be connected to wider theories of embodied behaviour? We aim to address these questions in future work.

Acknowledgements

Thanks to Andy Philippides and Lucas Wilkins for helpful discussions around the content of this work.

References

Ay, N., Bertschinger, N., Der, R., Güttler, F., and Olbrich, E. (2008). Predictive information and explorative behavior of autonomous robots. *The European Physical Journal B*, 63(3):329–339.

Ay, N. and Polani, D. (2008). Information flows in causal networks. *Advances in Complex Systems*, 11(01):17.

Barnett, L. (2009). Granger Causality and Transfer Entropy Are Equivalent for Gaussian Variables. *Physical Review Letters*, 103(23):1–10.

Belyakov, A. O., Seyranian, A. P., and Luongo, A. (2009). Dynamics of the pendulum with periodically varying length. *Physica D: Nonlinear Phenomena*, 238(16):1589–1597.

Bertschinger, N., Olbrich, E., Ay, N., and Jost, J. (2008). Autonomy: an information theoretic perspective. *Bio Systems*, 91(2):331–45.

Breiman, L. (1969). *Proability and Stochastic Processes: With a View Toward Applications*. Houghton Mifflin Company, Boston, MA.

Brooks, R. (1986). A robust layered control system for a mobile robot. *Robotics and Automation, IEEE Journal of*, (I):14–23.

Cuomo, K. and Oppenheim, A. (1993). Circuit implementation of synchronized chaos with applications to communications. *Physical Review Letters*, 71(1):65–68.

Evans, D. (2008). A computationally efficient estimator for mutual information. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 464(2093):1203–1215.

Janzing, D. and Balduzzi, D. (2012). Quantifying causal influences. *arXiv preprint arXiv:1203.6502*, pages 1–23.

Klyubin, A. S., Polani, D., and Nehaniv, C. L. (2008). Keep your options open: an information-based driving principle for sensorimotor systems. *PLoS ONE*, 3(12):e4018.

Kraskov, A., Stögbauer, H., and Grassberger, P. (2004). Estimating mutual information. *Physical Review E*, 69(6):16.

Lizier, J. T. and Prokopenko, M. (2010). Differentiating information transfer and causal effect. *The European Physical Journal B*, 73(4):605–615.

Lungarella, M. and Sporns, O. (2006). Mapping information flow in sensorimotor networks. *PLoS computational biology*, 2(10):e144.

Moioli, R. C., Vargas, P. A., and Husbands, P. (2012). Synchronisation effects on the behavioural performance and information dynamics of a simulated minimally cognitive robotic agent. *Biological cybernetics*, pages 407–427.

Pfeifer, R., Lungarella, M., and Iida, F. (2007a). Self-organization, embodiment, and biologically inspired robotics. *Science (New York, N.Y.)*, 318(5853):1088–93.

Pfeifer, R., Lungarella, M., Sporns, O., and Kuniyoshi, Y. (2007b). On the information theoretic implications of embodiment - principles and methods. In Lungarella, M., Iida, F., Bongard, J., and Pfeifer, R., editors, *50 Years of Artificial Intelligence*, volume 4850 of *Lecture Notes in Computer Science*, pages 76–86, Berlin / Heidelberg. Springer.

Pinsky, M. and Zevin, A. (1999). Oscillations of a pendulum with a periodically varying length and a model of swing. *International Journal of Non-Linear Mechanics*, 34(1):105–109.

Pitti, A., Lungarella, M., and Kuniyoshi, Y. (2009). Generating spatiotemporal joint torque patterns from dynamical synchronization of distributed pattern generators. *Frontiers in Neurorobotics*, 3(2).

Sauer, T. (2012). Numerical solution of stochastic differential equations in finance. In *Handbook of Computational Finance*.

Schmidt, N. M., Hoffmann, M., Nakajima, K., and Pfeifer, R. (2012). Bootstrapping Perception Using Information Theory: Case Studies in a Quadruped Robot Running on Different Grounds. *Advances in Complex Systems*, 16:1250078.

Schreiber, T. (2000). Measuring information transfer. *Physical Review Letters*, 85(2):461–464.

Staniek, M. and Lehnertz, K. (2008). Symbolic Transfer Entropy. *Physical Review Letters*, 100(15):1–4.

Thorniley, J. (2011). An improved transfer entropy method for establishing causal effects in synchronizing oscillators. In Lenaerts, T., Giacobini, M., Bersini, H., Bourgine, P., Dorigo, M., and Doursat, R., editors, *ECAL 2011: Proceedings of the Eleventh European Conference on the Synthesis and Simulation of Living Systems*. MIT Press.

Williams, P. and Beer, R. D. (2010). Information Dynamics of Evolved Agents. In *From Animals to Animats 11*, pages 38–49. Springer, Berlin / Heidelberg.

Zevin, A. and Filonenko, L. (2007). A qualitative investigation of the oscillations of a pendulum with a periodically varying length and a mathematical model of a swing. *Journal of Applied Mathematics and Mechanics*, 71(6):892–904.