Evolving Coordinated Behavior by Maximizing Information Structure – Olaf Sporns and Max Lungarella

O. Sporns and M. Lungarella, "Evolving coordinated behavior by maximizing information structure," presented at Tenth International Conference on Simulation and Synthesis of Living Systems (ALifeX), Bloomington, Indiana, USA, 2006.

Entropy Discussion Group
7/8/2006

Joseph Lizier
PhD student
The University of Sydney – Prof. Albert Zomaya
CSIRO ICT Centre – Dr. Mikhail Prokopenko

CSIRO

Evolving Coordinated Behavior by Maximizing Information Structure – Olaf Sporns and Max Lungarella

Abstract

Embodied systems actively structure information sampled by their sensors as they engage in sensorimotor interactions with their environment. Can information structure serve as an evolutionary principle that shapes behavior and leads to increased coordination? Here we address this question by attempting to evolve coordinated behavior in a simulated creature subjected to behavioral and information-theoretical cost functions. Our results show that maximizing information structure is highly effective in generating coordinated behavior, providing further support for a potential central role of actively generated information structure in embodied cognition.

CSIRO

Contents

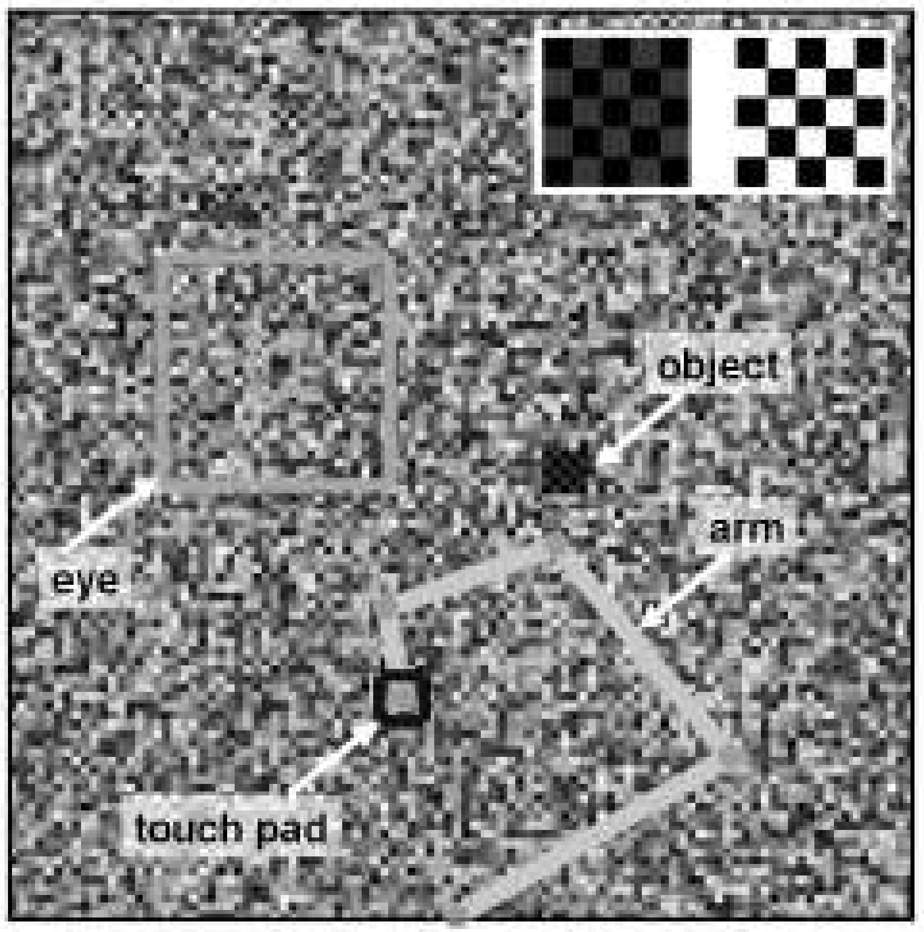
www.ict.csiro.au

CSIRO

Motivation: Can info structure serve as evolutionary principle?

- Importance of embodiement: “embodied systems are not passively exposed to sensory information but actively structure and exploit such information.”
- “*Information structuring* by sensorimotor activity and *information processing* by the nueral system are reciprically linked.”
- Previous work of the authors:
  - Quantitative evidence that info structure in sensory experience can be induced by effectively coordinated motor activity. Observed H↓, and I, Int, and C ↑.
  - Provides platform to investigate information driven evolution here.

Experimental setup



- 2D Environment:
  - Red, tactile object on noisy background
  - Object moves randomly unless touched
- Creature with sensors:
  - Eye
  - Touch pad on arm
- Creature has actuators:
  - Eye movement
  - Hand movement

Experimental setup – actuators

- Eye movement:
  1. Makes a saliency determination of each point in vision

$$S_{col} = f(\eta_R I_R + \eta_G I_G + \eta_B I_B + \eta_Y I_Y)$$

  2. Max of saliency map becomes target of vision movement.
  3. Jitter incorporated in determination of actual movement.

$$M_E = \varphi_E - \theta_E + \rho_E$$

- Touch pad movement:
  - Arm only moves if sum of saliencies in fovea exceeds a threshold ζ (then it moves towards foveal region ?).
  - Pre-training procedure allows creature to map hand position to angle configuration of it’s 4 joints.
  - Desired joint angles → movement required for each joint:

$$M_k = c_k(\varphi_k - \theta_k)$$

  - Jitter added only to the fourth joint angle

CSIRO

Experimental setup – genome and GA

- Genome consists of 11 alleles:
  - Four saliency factors η_R, η_G, η_B and η_Y
  - Amplitude of random jitter in the eye ρ_E
  - Scale factors for four arm joints c_k
  - Threshold for arm movement ζ
  - Amplitude of random jitter in the last arm joint ρ_4
- Alleles mutated at each generation, with a decreasing mutation "rate".
- 60 generations used
  - 10 individuals per generation. Individuals evaluated in 250 time steps.
  - Single individual selected for given cost function (see next slide) from each generation for mutation.

Experimental setup – Cost functions and performance measures

- Variety of performance measures used as cost functions used; all performance measures monitored for every cost function.
- Based on sensory data from the eye's fovea and touch pad. (Two 5x5 time series).

Behavioural (task based)

Information theoretic

Control

Cost Func.		*Description*
B	*foveation*	maximizing time for which distance between eye and object is less than 2.5 pixels
	touch	maximizing time for which object is touched
	fovtouch	conjunction of *foveation* and *touch*
	maxred	maximizing color red in the fovea (Eq. 7)
I	*negH*	minimizing entropy
	MI	maximizing mutual information
	Intg	maximizing integration (Eq. 8)
	Cplx	maximizing complexity (Eq. 9)
C	*H*	maximizing entropy
	negCplx	minimizing complexity (Eq. 9)

Information theoretical cost functions

- Integration – “the amount of statistical dependence within a system X of elements x_i”:

$$I(X) = \sum_i H(x_i) - H(X)$$

  - Sporns: “A multivariate extension of Mutual Info”

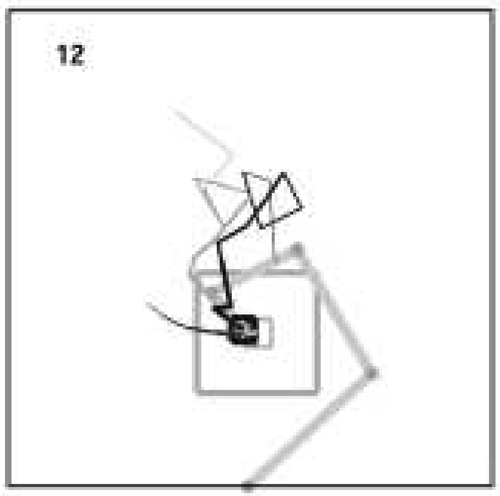
- Complexity – “expresses the degree to which globally integrated and locally segregated information coexists in a data set”:

$$C(X) = H(X) - \sum_i H(x_i \mid X - x_i)$$

  - Ordered: H(x) low, $H(x_i|X\text{-}x_i)$ low → C(X) low
  - Chaotic: H(x) high, $H(x_i|X\text{-}x_i)$ high → C(X) low
  - Complex: H(x) med-high, $H(x_i|X\text{-}x_i)$ low → C(X) high
  - Strong parallels to Crutchfield’s Transient Information

CSIRO

Results – Evolving for complexity



- Tracking of object observed.
- Arm tended to track eye as it scanned environment.
- Exploration occurred even after object was acquired (small amplitude visual and tactile scanning movements).
- Complexity ↑ with generations.
- Importantly, behavioural measures ↑ with generations also.
- Evolution for information structure obviously has some utility.

CSIRO

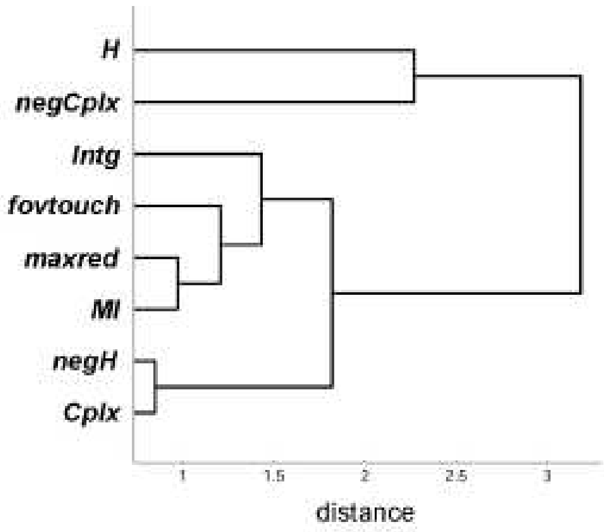
Results – all cost functions / performance measures

Construct	*Cost Function*	*Performance Measure*					
		H(X)	MI(X)	I(X)	C(X)	K_{red}	$\tau_{E,J}$
random		3.52 (0.63)	0.04 (0.04)	23.77 (18.17)	0.30 (0.29)	0.20 (0.28)	0.05 (0.14)
	negH	1.02 (0.07)	0.17 (0.01)	96.16 (11.40)	1.49 (0.07)	0.80 (0.03)	0.68 (0.03)
	MI	1.77 (0.19)	0.34 (0.04)	111.06 (9.63)	0.69 (0.08)	0.77 (0.02)	0.72 (0.08)
	Intg	1.40 (0.12)	0.28 (0.04)	120.56 (4.32)	1.10 (0.14)	0.79 (0.02)	0.71 (0.04)
	Cplx	1.25 (0.23)	0.17 (0.02)	92.24 (7.77)			0.64 (0.03)
	maxred	1.38 (0.23)	0.30 (0.08)	122.92 (5.37)			0.77 (0.07)
	fovtouch	2.02 (0.30)	0.33 (0.04)	106.88 (15.13)	0.67 (0.17)	0.77 (0.02)	0.79 (0.06)
	H	3.65 (0.19)	0.02 (0.01)	13.40 (1.19)	0.16 (0.01)	0.00 (0.00)	0.00 (0.00)
nojitter		1.94 (0.21)	0.24 (0.06)	85.68 (10.74)	0.70 (0.03)	0.74 (0.12)	0.56 (0.17)
ideal		2.01 (0.05)	0.31 (0.01)	79.46 (2.64)	0.47 (0.02)	0.75 (0.03)	0.96 (0.01)

- Evolving for complexity or max red also results in maximisation of the other metric.

Results – similarity between cost functions

- Information theoretic cost functions evolved:
  - Preference for the colour red.
  - Increased moving amplitude of the arm (scaling factors for joints)
  - Eliminated jitter of eye
  - Lowered threshold for arm activation
- Other similatiries:
  - MI and maxred
  - H and negCplx – little diff to random
  - Cplx and negH



Summary

- Maximising information structure appears highly effective in producing coordinated behaviour here.
- negH, MI, Int, Cplx all lead to similar coordinated behaviour solutions.
- Evolution for complexity was subtly different – continued scanning generated additional information structure. Importance of embodiment.
- Can extend beyond info from sensory system to info from neural representations: “neural architectures may only become fully realized if their neural complexity is matched by rich and highly structured sensory data”.

CSIRO

Complexity vs Transient Information

Complexity

- Looking at convergence of joint entropy of a multivariate to total joint entropy value as subset size increases.

Transient Information

- Looking at convergence of block entropy of a 1D variable to it's asymptote as block length increases.

$$C_N(X) = \sum_{k=1}^{n} [(k/n)I(X) - \langle I(X_j^k)\rangle].$$

$$C_N(X) = \sum_{k=1}^{n} [\langle H(X_j^k)\rangle - (k/n)H(X)].$$

$$\mathbf{T} = \sum_{L=1}^{\infty} L\left[h_\mu(L) - h_\mu\right]$$

$$C_N(X) = \sum_{k=1}^{n/2} \langle MI(X_j^k; X - X_j^k)\rangle.$$

- Similar, but are technically measuring slightly different things.

CSIRO